# UPPER SECONDARY SCHOOL STUDENTS' PERCEPTIONS OF USING MOBILE PHONE APPLICATIONS IN THE FORMAL ASSESSMENT OF SPEAKING SKILLS

Minor's thesis
Katja Jääskeläinen

**JYVÄSKYLÄN YLIOPISTO**

| Tiedekunta – Faculty | Laitos – Department |
|---|---|
| Humanistis-yhteiskuntatieteellinen tiedekunta | Kieli- ja viestintätieteiden laitos |

| Tekijä – Author |
|---|
| Katja Jääskeläinen |

| Työn nimi – Title |
|---|
| UPPER SECONDARY SCHOOL STUDENTS' PERCEPTIONS OF USING MOBILE PHONE APPLICATIONS IN THE FORMAL ASSESSMENT OF SPEAKING SKILLS |

| Oppiaine – Subject | Työn laji – Level |
|---|---|
| Englanti | Sivututkielma |

| Aika – Month and year | Sivumäärä – Number of pages |
|---|---|
| Huhtikuu 2020 | 49 + 1 liite |

Tiivistelmä – Abstract

Tutkimuksen aiheena on lukion oppilaiden käsitykset älypuhelinsovellusten käytöstä englannin suullisen kielitaidon arvioinnissa. Suullinen kielitaito, sen arviointi ja mobiilioppiminen (Mobile assisted language learning, MALL) ovat 2000-luvun keskeisiä teemoja. Tutkimuksen tarkoituksena on yhdistää nämä kolme aihealuetta ja osoittaa, että älypuhelinsovellukset voivat monipuolisuutensa ansiosta olla hyödyllisiä suullisen kielitaidon arvioinnissa, koska aiemmat tutkimukset ovat tutkineet MALL:ia vain kielitaidon eri osa-alueiden opettamisen ja oppimisen näkökulmista. Tutkimukseen osallistui 33 oppilasta itäsuomalaisesta lukiosta. Tässä empiirisessä tutkimuksessa käytettiin kvantitatiivisia ja kvalitatiivisia tutkimusmenetelmiä. Aineisto kerättiin Webropol-kyselylomakkeella. Tulosten mukaan mieluisimpia koemuotoja olivat keskustelut, ääneenlukutehtävät ja puheet/esitelmät. Suurin osa vastaajista halusi tehdä kokeet parin ja osa opettajan kanssa. Luokan edessä tekeminen jakoi mielipiteitä. Suurin osa halusi sanallisen tai sekä sanallisen että numeroarvioinnin. Valmiista vaihtoehdoista valittuina suosituimpia älypuhelinsovelluksia olivat puheen äänittämissovellukset; draamasovellukset, kuten elokuvientekosovellukset; ja roolihahmo- ja avatar-sovellukset, kuten nukketeatteri- ja kertomussovellukset. Sovellusten hyötyinä nähtiin mm. jännityksen väheneminen, uudet mahdollisuudet, luovuus ja helppous sekä oppilaalle että opettajalle. Haittoina mainittiin mm. keskittymisongelmat älypuhelinta käytettäessä, huijausmahdollisuudet, tekniset ongelmat ja teknisen osaamisen taso sekä oppilailla että opettajilla. Johtopäätöksenä voidaan todeta, että lukion oppilaat olivat ainakin osittain valmiita ottamaan älypuhelinsovellukset avuksi suullisen kielitaidon arviointiprosessiin ja että he näkivät ne mahdollisuutena, mutta myös epävarmuutta niiden toimivuuden ja oman osaamisen suhteen on. Jatkotutkimusta tästä aiheesta olisi hyvä tehdä isommalla otoksella ja myös opettajien käsityksistä, osaamisen tasosta ja koulutustarpeista.

| Asiasanat – Keywords | speaking skills, assessing, MALL, survey/questionnaire |
|---|---|

| Säilytyspaikka – Depository | JYX |
|---|---|

| Muita tietoja – Additional information |
|---|

Table of contents

# 1 Introduction

The general research areas of the present study are speaking skills in the context of English as a Foreign Language (EFL), the assessment of speaking skills and Mobile assisted language learning (MALL). They are all relevant and important areas in the 21st century. Previous research in this field has concentrated on teaching and learning language skills in English by the means of MALL. They have disregarded assessment and often also speaking skills while other skills have been taken more into consideration. In the previous studies in the area of MALL, Abugohar, Yunus and Rashid (2019), Cochrane (2015) and Fučeková (2018) have emphasized the importance and the possible effectiveness of using smartphone applications in teaching and learning language skills in the EFL context in the 21st century and that smartphone applications have to be taken into consideration since they have become an inseparable part of learning process. All in all, the use of MALL in the world we are living in is obvious.

The importance of speaking skills has been acknowledged in the basics of the new national curriculum for upper secondary school (LOPS2019), in the development of matriculation examination, and in the Common European Framework of Reference (CEFR). Thus, more and more emphasis is being given to speaking skills. The basics of LOPS2019 were published in November 2019 and it will be in use in 2021. One of the basic aims in foreign language studies is to encourage students to use different languages in various manners and to take into consideration all the sections of language skills. Learning languages is seen as an important tool to be used and to influence in the society and in the international world. Speaking skills are divided into three categories: acting in interaction, interpreting *texts* and producing *texts* (LOPS2019: 129). Tergujeff and Kautonen (2017:13-14) state that speaking skills are not defined in very detailed way in LOPS. They state that the term written and oral *texts* can be misleading when reading the present LOPS2015 since it is so easy to think about written *text* when the term *text* is used.

According to LOPS2019, in the assessment of language skills various kind of tools should be used and feedback should be encouraging. Students can have a separate diploma after having done the optional course ENA8 *Speak: Communicate and influence* (*Viesti ja vaikuta puhuen*) and its separate exam. As Tergujeff and Kautonen (2017:16-17) emphasize, an exam of speaking skills is also being planned to be a new part of the matriculation examination and therefore teaching and learning speaking skills should be as important as learning other language skills. The assessment of speaking skills can provide a lot of questions: What are

speaking skills? What is important to learn and assess when considering speaking skills? How to have time to assess these skills in a reliable and valid way? How students would like to be assessed? MALL could be used to help teachers in the assessment of speaking skills.

The aim of the study was to investigate how students perceive the use of smartphones in the formal assessment of speaking skills and how they want to be assessed in speaking skills if they were able to make decisions on the assessment. The research questions in this empiric study were the following: 1) *How would students like to be assessed formally in speaking skills?*, 2) *How would students like to be assessed formally in speaking skills when using smartphone applications?* and 3) *Which applications that they already know could be used in assessing speaking skills in their opinion?* To get answers to these research questions a questionnaire in Webropol was used to collect data and thus both quantitative and qualitative data was received.

In addition to answering the research questions, this study provides suggestions for teachers on how to create the situations in the formal assessment of speaking skills as natural and easy-going as possible for the students by the means of mobile phone applications. For example, students who do not want to record their voice for any purpose could use drama applications and be in a role and thus their character or avatar could protect them, and exams could be done safely at home. Moreover, there are practical problems which mobile phone applications can solve since the exams can be made at home. For example, there will no more be need for to invigilate students in one classroom when a teacher is in another classroom with the student having the test when teaching in lower or upper comprehensive schools or to find a quiet place without any distractions to provide similar circumstances to all students and thus secure reliability. Furthermore, one of the aims of this study was to utilize students' MALL-knowledge and find new interesting, practical, easy to use and authentic, i.e. real-life applications to use in the formal assessment of speaking skills.

This thesis is organized into six chapters. First, speaking skills and assessing will be looked at in Chapter 2 and definitions of speaking skills, its importance and nature are presented. Moreover, it takes a closer look to test usefulness, the types of assessment and the assessment of speaking skills. The end of Chapter 2 will be concentrating on MALL and the advantages of it. Chapter 3 presents the participants, questionnaire and methods of analysis, while chapter 4 introduces the results and their analysis. Chapter 4.1 introduces background information of the respondents and their previous formal assessment of speaking skills, while in chapter 4.2 respondents' suggestions of different types of exams and the use of smartphone applications will be looked at. In chapter 4.3 SWOT analysis will be used when taking a closer

look to respondents' perceptions of the use of smartphone applications in the assessment of speaking skills. Finally, Chapters 5 and 6 conclude the study – in Chapter 5 there is a discussion of the results and in Chapter 6 there are the conclusions and suggestions for further studies.

## 2 Background theory and framework

Since the aim of this study is to comprehend the concepts of speaking skills, assessment and MALL, and to combine these three concepts in an efficient way in the context of EFL, all these concepts will be introduced in this chapter. First, in chapter 2.1, I will describe and discuss language skills, concentrating on speaking skills and its importance in teaching and learning English. Moreover, I will explain which speaking skills could be seen as the most important ones. In chapter 2.2, I will present and discuss test usefulness, different types of assessment and the assessment of speaking skills. In chapter 2.3, I will discuss the advantages of mobile phone applications and therefore present studies on the field of MALL concerning teaching and learning of language skills.

**2.1 Speaking skills**

2.1.1 Definitions of speaking skills

Speaking skills and its subcategories have been defined in various ways. Tergujeff and Kautonen (2017: 12-21) conclude that the terms language skills and speaking skills can be misleading since there are always other skills and knowledge involved when using a language. Canale and Swain (1980) divided communicative competence into 1) linguistic, 2) sociolinguistic and 3) strategic competences. In their division, linguistic competence consists of the knowledge of the language code such as grammatical rules, vocabulary and pronunciation. Sociolinguistic competence means the knowledge of the socio-cultural code of language use such as appropriate vocabulary, register, style and politeness. Strategic competence consists of verbal and non-verbal skills which enable us to overcome difficulties in communication situations. Bachman and Palmer (Bachman 1990, as quoted by Bachman and Palmer 1996: 67-75) have divided language ability into 1) language knowledge and 2) strategic competence i.e. a set of metacognitive strategies. They state that the combination of language knowledge and metacognitive strategies provides language users their ability to create and

interpret discourse. According to them, language knowledge includes organizational and pragmatic knowledge. Organizational knowledge includes grammatical knowledge, knowledge of vocabulary, textual knowledge and knowledge of cohesion while pragmatic knowledge includes functional knowledge and sociolinguistic knowledge. The second category of language ability, strategic competence, includes goal setting, assessment and planning. Bachman and Palmer discuss also metacognitive strategies in language use and language test performance. All in all, as we can see, speaking skills are not a simple issue to define and the terms knowledge, competence, strategy and abilities are used as having same meanings, as are also the terms strategic and meta. Moreover, Canale and Swain and Bachman and Palmer have placed the strategic competence into different categories even though they all have it in their categories of competences, knowledges or abilities.

CEFR guides teaching and assessing speaking skills. In the improved CEFR communicative language competences (Council of Europe 2018: 130-142) are divided into three categories: 1) linguistic, 2) sociolinguistic and 3) pragmatic competences. In the assessment scales of CEFR linguistic competences consist of six subcategories: general linguistic range, vocabulary range and control, grammatical accuracy, orthographic control and phonological control. The new modified division of Phonological control consists of overall phonological control, sound articulation and prosodic features – and it will be discussed more in Chapter 2.2.3 *Assessing speaking skills*. In CEFR sociolinguistic competences are defined as sociolinguistic appropriateness, while pragmatic competences are divided into the following six subcategories: flexibility, turntaking, thematic development, coherence and cohesion, propositional precision and spoken fluency.

Moreover, speaking skills can be divided into routine skills and improvisational skills (Bygate 1987 and Huhta 2010, as quoted by Ahola 2017: 156-158). Ahola states that in exams students may use routine skills when there are tasks that demand common phrases or structures, but they should also be prepared to unexpected issues in interaction and be able to use improvisational skills. She mentions that strategic competence helps examinees in these kind of unexpected situations in interaction since it may compensate their linguistic competence. Ahola underlines that both routine and improvisational skills should be assessed.

Tergujeff and Kautonen (2017: 12-21, 170-171) state that pronunciation is an important part of speaking skills and that it can be defined in a narrow way as producing singular phonemes or in a broader way in which it contains both phonemes and prosody. Kuronen (2017: 59) defines that prosody consists of intonation, rhythm as the variation of unaccented and accented syllables, and facilitations in pronunciation, e.g. reduction and assimilation. Kuronen

states that the division into prosody and phonemes is important when learning a language since it helps focus on different learning purposes but that otherwise this division is not absolutely clear phonetically. Tergujeff and Kautonen (2017: 170) give an example of accented syllables and their influence on meaning: in English nouns the stress is on the first syllable and in verbs on the second, e.g. REcord vs. ReCORD. In phrases and clauses, the stress is usually on words which have most of the meaning. All in all, Tergujeff (2017: 170-171) describes prosody and phonemes to be an important part of speaking skills when considering the aim to be understood in communication. For instance, when learning English, intonation, certain phonemes which differ from Finnish and the distinction of certain phoneme pairs is crucial for understanding, e.g. the voiceless and voiced minimal pairs p/b, k/g, t/d, v/f and s/z.

2.1.2 Importance and nature

Tergujeff and Kautonen (2017: 12-21) discuss the importance of speaking skills and the fact that when learning your mother tongue, you learn first to speak it and only after that to write it. However, in the EFL classroom context it is often the case that speaking skills are learned after writing skills have been taught. They state that other areas – e.g. grammar, vocabulary and writing - are often considered more important than speaking skills. Moreover, Tergujeff and Kautonen discuss the importance of speaking skills in everyday and working life. They state that it should not be taken for granted that speaking skills would be learned automatically in free time when hearing and using a foreign language. Therefore, it is important to teach and assess speaking skills in EFL lessons.

Johnson (2013: 278-299) discusses the four language skills which are divided into productive ones – writing and speaking – and receptive ones – reading and listening. Bachman and Palmer (1996: 75-76) critique this approach to divide language skills in terms of channel (audio, visual) and mode (productive, receptive). According to them, this division was very influential in language testing during the second half of the 20th century. The disadvantages in this division are that divergent language use tasks are classified under a single 'skill' and that it is not taking into consideration the fact that language use happens in a particular setting, not in a vacuum. They give an example of different settings, one being a face-to-face conversation and another just listening to a radio newscast. They both involve listening, but the activities and settings are different. In their opinion, it is not useful to think in terms of 'skills' but more broadly, as a concept of 'ability-task' which means that specific activities and tasks in which language is used are more important to take into consideration than separate 'skills'.

Speaking has its own special features when compared with the other three skills. Council of Europe (2020) has defined five descriptors for spoken language in the CEFR: range, accuracy, fluency, interaction and coherence. In interaction non-verbal and intonational cues and turntaking are mentioned which are special features for speaking. Luoma (2004: 10-11) mentions communicative effectiveness and comprehensibility which are also important in interaction. Luoma (2004: 27-28) states that in applied linguistics, the ability to speak a language can be seen as a meaningful interaction with social and situational features and needs in it and that in this interaction, either form or meaning can be emphasized.

**2.2 Assessing**

2.2.1 Test usefulness

Bachman and Palmer (1996: 66-67) state that when language skills are defined for purposes of assessing and measurement, we are defining a construct. Construct must be defined precisely enough in a way that other characteristics which could affect the test results would not affect it. It should also be defined separately for every test situation according to its purpose, examiners and target language use (TLU) domain. For example, politeness markers in each situation should be taken into consideration when assessing.

Bachman and Palmer (1996: 17-40) discuss test usefulness in their book *Language Teaching in Practice.* Test usefulness is about qualities of language tests and it is the most important issue when an exam is designed, developed and used (Bachman and Palmer 1996: 17-40). They propose a model of test usefulness which includes six test qualities: reliability, construct validity, authenticity, interactiveness, impact and practicality. Bachman and Palmer argue that the complementarity and the appropriate balance of these six qualities is to be achieved in given test situations – instead of emphasizing the tension between them or their minimum or maximum acceptable levels.

Now I will discuss further these six test qualities of test usefulness using Bachman and Palmer (1996:17-40) as my primary source of information. The first five qualities – reliability, construct validity, authenticity, interactiveness and impact – are about the use of test scores. Conversely, the sixth quality, practicality, is about the ways the test is used or if it is used at all. They state that reliability means that measurement is consistent and that there is no variation in test scores which could be caused by other factors than the construct itself. Therefore, tasks should be carefully designed. Construct validity means that we want to have results that reflect

the certain language skill or ability which we are supposed to be measuring. Authenticity means that a test is authentic in a way that the skills used in the test could be useful in real-life situations. We use language in unique situations, i.e. language use domains. Target language use (TLU) domains can be divided into two general types: real-life and language instruction domains (Bachman and Palmer 1996: 43-45). In real-life domains language is used for communication and in language instruction domains it is used for teaching and learning of it. Interactiveness measures test takers' involvement of their language ability and knowledge and tells about the degree to which the assessed constructs are used when doing the task. Impact means the ways in which the test results affect individuals and their future placements, education and the whole society. Another term to describe these effects is washback. Practicality means the resources which we have or have not in the process of design, development and use of a test. For example, time, place, teachers and equipment are resources. All in all, Bachman and Palmer state that test usefulness is a useful tool on the field of language testing since it helps us to understand and see the interaction between these six qualities in specific testing situations.

Bachman and Palmer (1996: 18-19) state that in a classroom test, teachers may want to have higher levels of authenticity, interactiveness and impact. Conversely, they state that in large-scale tests, involving important decisions about students' future placements, the aim is to have the highest possible levels of reliability and validity. Therefore, they state that reliability and validity can be referred to as essential measurement qualities because they provide the major justification when test scores are used as a basis of decisions.

Furthermore, Luoma (2004: 170-191) discusses reliable and valid speaking assessment. She (2004: 28) underlines construct validity in exercises where there are more than one person speaking. This situation can be possible when using applications. She points out that since speaking is interactive it should be taken into consideration when planning rating criteria. Examinees should know how to work interactively with other speakers in exams. She emphasizes that it is important that developers of speaking assessments understand that speaking is about both construct and context. According to her, it is also important to consider construct, tasks and rating criteria and also remember to inform examinees about them. Ahola (2017: 155-156) discusses assessment of speaking and its validity, reliability, impact and ethic. She states that these are also important issues to consider when developing exams in speaking skills. According to the concept of ethic, assessment should be fair, consistent and supporting, not discouraging. Ahola states that assessment can affect learners' impressions of themselves as L2 users.

The importance of my study related to test usefulness can be seen for example when considering practicality. When using smartphones in assessing speaking skills, there are advantages on practicality since it is less time consuming and there is no need for an empty space or classroom when having the test. Moreover, other resources like teachers or other personnel are not needed to invigilate the examination situation when tests are made at home by using students' own smartphones.

2.2.2 Assessment types

Luoma (2014:190-191) indicates that current research in assessing speaking has been done more on formal assessment of speaking skills, i.e. on formal proficiency tests. But she states that since learning-related and informal assessments are increasing, more studies should be conducted in this informal context of assessment. However, in my study I concentrate on formal assessment since I have experienced it to be more demanding in teachers' work.

Keurulainen (2013: 37-45) states that since the idea of learning and learners has changed from a behaviorist one towards a constructivist one, the concept of assessment has changed too. He describes that behaviorism is about quantity and quality of results and it is controlling what have been learned while the basic idea in constructivism is to scaffold the learning process. Clearly, the work of a teacher is nowadays more to be an instructor for the learners. Keurulainen indicates that in assessment, the behaviorist way of thinking is not enough alone. The assessment need to be an ongoing and active part of the process of teaching and learning. The change has been happening from a declaratory assessment which is done from the outside of the learner towards to a developing assessment (Hakkarainen 1980, as quoted by Keurulainen 2013: 37-45).

Keurulainen explains that the concept of assessment has been changed and broadened in several ways. Firstly, the only goal of assessment is no longer to give grades, and secondly, those who give feedback or assess are no longer only teachers but also students acting as self-, peer- or group assessment givers (Anttila 2011, as quoted by Keurulainen 2013: 37-45). Keurulainen (2013: 37-45) discusses the concept of assessment when he explains the difference between formative and summative assessment (Michael Scriven 1967, as quoted by Keurulainen 2013: 37-45). This division can be made according the time during which the assessment is done. Formative assessment happens during the process of learning and summative assessment is done at the end of this process. Formative assessment deals with issues

that emerge during the learning process and its function is therefore reactive and it gives direction to the learning process.

Furthermore, Keurulainen (2013: 37-45) mentions an important difference: assessment *of* learning and assessment *for* learning. Assessment *of* learning is the traditional way of seeing assessment as assessing what has been learned and giving a grade while assessment *for* learning is assessment which helps the learner to know which way to go in one's learning process, i.e. it guides the learner to the right direction. Assessment should give an opportunity to learn and develop (Ecclestone 2012, as quoted by Keurulainen 2013: 37-45). Formative assessment should be used since learning results are improving when using it (Leahy and Wiliam 2012, as quoted by Keurulainen 2013: 37-45). All in all, Keurulainen states that the function of assessment can be normative in which learners are compared and reliability is important, or it can be based on criteria made beforehand.

2.2.3 Assessing speaking skills

After having introduced the different ways to categorize speaking skills and discussed speaking skills and competences in Chapter 2.1.1, I will now focus on their importance in assessing as there are several issues that should be taken into consideration when assessing speaking skills.

Tergujeff and Kautonen (2017: 12-21) emphasize that it is important to assess speaking skills. Luoma (2004: 1-11) also considers that speaking skills are an important object of assessment and that assessing speaking skills is a challenging task since there are so many factors which influence our impression of an examinee's speaking skills. Kuronen (2017: 68-70) also points out that prosodic features are factors which influence decisions made in the assessment of speaking skills. Moreover, Luoma (2004: 29-35) argues that we suppose that test scores would be accurate and just for our purposes. She states that pronunciation accuracy is easy to choose as a construct since it can quite easily be judged against a norm. Luoma (2004: 29-35) states that the language is different in different contexts and with different purposes and that this affects the developing and design of speaking tasks. Moreover, she notes that the assessment formats can be various since there can be individual, pair and group tasks. Bachman and Palmer (1996: 75-76) also state that when designing language tests, we should take into consideration task characteristics as setting, input, expected response, and relationship between input and response, and the areas of language ability and topical knowledge.

Luoma (2004: 44-45) discusses the differences between two test modes: live i.e. face-to-face interaction or tape-based modes, while there is no discussion about videos. Luoma

considers that test modes can be divided according to the construct. She claims that the construct can be clearly more about spoken *interaction* (live) or spoken *production* (tape-based). In my opinion, the construct of a live situation can also be production and the construct of a video situation can be interaction because a live situation can be a presentation and a video situation can be made using new technologies and MALL in pairs or in groups, e.g. demonstrating a TLU situation in a shop. She states that the live test mode is two-directional since speakers' reactions are involved. But in my opinion, a live or face-to-face test can be a presentation and therefore a one-directional task. Even though she mentions the possibility of video teleconferencing when discussing live possibilities, it is only for geographical reasons. She also mentions the possibility to use a phone but not more about it.

In CEFR 2001 there were weaknesses in the existing phonology scale and new scales were created after the weaknesses had been identified (Piccardo 2016: 9-11). The new analytic grid replaced the original CEFR scale for Phonological Control (Piccardo 2016: 23). According to Piccardo (2016: 23), the aim was to provide an efficient and easy to use tool for assessing work and to help learners to know better what they are expected to learn. There were 34 descriptors for Phonological control which were finally grouped into three categories in the new phonology scale: general phonology scale, pronunciation (sound articulation) and prosody (intonation, stress and rhythm) (Piccardo 2016: 15-20). Piccardo states that the core areas during the process of creating these three descriptors were the following five: articulation (including pronunciation of phonemes), prosody (including intonation, rhythm and word/sentence stress, and speech rate/chunking), accentedness (accent and deviation from a 'norm'), intelligibility (i.e. actual understanding of an utterance by a listener) and comprehensibility (i.e. listener's perceived difficulty in understanding an utterance). She emphasizes the importance of the relation between these core areas and that native-like pronunciation and accentedness should not be the focus in language teaching and learning since intelligibility is much more important. Thus, in the new CEFR even in the level of C2 accentedness is no longer a requirement.

Piccardo underlines (2018: 14) that there is a gap between the research and teaching when concerning the assessment of pronunciation and that teachers should have more formal training, assessment tools and research-informed support since they are "often left alone and very often neglect teaching of pronunciation, thus disadvantaging precisely those learners that would mostly benefit from such instruction." CEFR can be used as a tool of assessment which is based on research on the field of education. Piccardo (2018: 13-14) states that the two major factors which are likely to have a negative impact on pronunciation training are its quality and

difficulties in the assessment of pronunciation. She also states that raters' knowledge of phonology, clear assessment criteria and the construct validity are important, and that phonological features should not be mixed up with linguistic features like grammatical or lexical control.

As it can also be seen in CEFR (Council of Europe 2018), the spectrum of speaking skills and competences is extremely wide, and teachers are those who decide what kind of exams they have and which competences and descriptors they emphasize. It is important to have various kind of exams on speaking skills and learn how to assess them properly. MALL can be a useful tool for teachers. Moreover, scales like CEFR can be useful tools and curriculum need to be taken into consideration in the process of assessing speaking skills.

The improved CEFR presents two main task categories for the assessment of the spoken production (Council of Europe 2018: 68-73): sustained monologue and addressing audiences. The descriptor scales in sustained monologue are describing experience, giving information, public announcements and presenting your argument, e.g. in a debate. Addressing audiences involves presentations and speeches. Interaction activities in the CEFR are the following (Council of Europe 2018: 83-92): understanding an interlocutor, conversation, informal and formal discussion, goal-oriented co-operation, obtaining goods and services, information exchange, using telecommunications and interviewing and being interviewed. MALL could be used in all these activity types, but it is not at least explicitly mentioned in CEFR. Interaction strategies according to CEFR (Council of Europe 2018: 100-102) are turn taking, cooperating and asking for clarification (scales from pre-A1 to C2). Mediation is one concept in CEFR (Council of Europe 2018: 103-105): "in mediation the user/learner acts as a social agent who creates bridges and helps to construct or convey meaning" while language is used as a communication tool, collaborating, encouraging and passing information. Mediation strategies (Council of Europe 2018: 126-127) are divided into two categories: 1) Strategies to explain a new concept (linking to new knowledge, adapting language, breaking down complicated information) and 2) strategies to simplify a text (amplify a dense text, streamlining a text).

**2.3 Mobile Assisted Language Learning (MALL)**

2.3.1 Advantages when used in learning and assessing

Tergujeff and Kautonen (2017: 12-21) state that having exams on speaking skills has been seen as a time-consuming process since it takes time to organize and evaluate them. This is one of

the reasons why I want to know and learn more about the possibility to use smartphones in the formal assessment of speaking skills. Smartphones provide us an easy and effective possibility to record speech and in addition to make videos and movies in a way that a product or a task can be listened and watched several times. In a similar way, smartphone applications can be used in teaching and learning speaking skills. Moreover, Ahola (2017: 153-168) discusses the use of mobile phones when assessing speaking skills. She also states that it is more valid when mobile phones or tablets are used to record since the product can be listened to several times and thus the teacher can have a better understanding of students' speaking skills. Furthermore, when using students' own smartphones or tablets the formality of the exam can be reduced. It is easy to record speech since there are microphone, video camera and speakers in smartphones and tablets. Moreover, there are several applications with which pictures, drawings and videos can be added to the speech, e.g. *QuickVoice, iMovie* and *Animation Creator HD* (Ahola 2017: 153-168). In addition, there are applications like *Morfo* in which you can use a character when speaking English. Mobile phones are mobile which is an advantage since formal proficiency tests can be done wherever it would be the most comfortable place to do it for the student. This can also reduce tension in speaking skills exams.

Kuronen (2017: 59-72) emphasizes that pronunciation should be taught explicitly by explaining and illustrating the phonetic features in it. Phonetic awareness of phonemes and prosody is important. Mobile phones can help nowadays' learners to observe themselves and thus we do not need language laboratories anymore to do this. Learners can easily record their own pronunciation and listen to and make observations of it. Perception helps them to improve their pronunciation. The importance of repetition and feedback should not be forgotten (Kjellin 2002, as quoted by Kuronen 2017: 59-72). For example, there is a useful speech analyzing program *Praat* which can be used to learn prosodic features (Boersma and Weenink 2016, as quoted by Kuronen 2017: 59-72). This program is helpful since it gives visual feedback to learners. The only disadvantage of *Praat* is that it is still quite expensive. Moreover, Derwing and Munro (2015: 126) emphasize visual representation of speech and speech technology such as *Praat*. They define *Praat* to be a sound spectrograph giving visual speech analysis. Derwing and Munro (2015: 127) mention also *Visipitch* by Kay Elemetrics which is a pitch track. They consider it to be useful since pitch is a relatively straightforward concept and thus it is quite easy for learners to interpret the visual patterns, such as pitch rises and falls (Chun et al. 2008, as cited in Derwing and Munro 2015:127).

Kuronen (2017: 68-70) indicates also that an acoustic vocal map can be useful when visualizing in learning to speak and that MALL is a useful tool for self-assessment. Moreover,

he mentions online materials, such as pronunciation pages, applications and image dictionaries. In addition to all his suggestions, I could suggest *Sounds of speech* on the Internet and as an application. Kuronen underlines that both phonemes and prosody are important to be learned, as are the perception and the articulation.

In my opinion, when videos are made, they provide a great opportunity to observe one's own and others' pronunciation and other skills that are needed in communication. They offer an opportunity to self, peer, group and teacher assessment. Making videos or vlogs is a process of learning in which assessment *for* learning and formative assessment can be used. For example, *Imovie, Puppet Pals* and *Morfo,* can be useful applications when making videos and vlogs. Tergujeff, Heinonen, Ilola, Salo and Kara (2017: 102-104) also discuss drama pedagogy and MALL applications, e.g. *Puppet Pals* and *Sock Puppets*. They also mention *QuickVoice* application with which learners can record their speech and get personal feedback in a very detailed way. For example, places in which the pronunciation systematically differs from the pronunciation of the language being learned are underlined. Furthermore, learning environments on the Internet are mentioned and learners' possibility to use their mobile phones which are a great benefit for rehearsing speaking skills. Learning need not only to occur in face-to-face situations anymore, and homework can be done by using technology. In my opinion, it is important to remember that all that can be used in learning can also be used in assessing.

Since there can be same applications in Computer Assisted Language Learning (CALL) and MALL, also some technology of CALL is presented next. Derwing and Munro emphasize the use of technology in L2 pronunciation instruction (2015: 128-130). They use a term Computer-assisted pronunciation training (CAPT) which should not be mixed up with automatic speech recognition (ASR) being one part of CAPT. They propose a software called *Second life* in which there is a virtual world since according to them virtual worlds offer exposure to spoken language which is important since for some students it is the only way to be exposed to spoken L2. As Kuronen also emphasizes, perception is important when learning speaking skills**,** so do Derwing and Munro. In this study smartphone applications can provide virtual worlds as assessing contexts for students. Moreover, assessment situations can also be seen learning situations. Derwing and Munro (2015: 130) state that there are several advantages of technology for pronunciation teaching and learning. They mention the strengthened perception skills, the extremely interesting contents and opportunities to interact orally in L2 when various methods are used. According to them, technology can also facilitate teachers' work. But they emphasize that technology should be seen one tool among others and that there can also be disadvantages. Moreover, they underline that it is important to understand "the

foundations of pronunciation research and pedagogical knowledge to exploit the benefits that technology has to offer" (Derwing and Munro 2015: 130).

2.3.2 Studies of MALL

In the following three studies of Abugohar, Yunus and Rashid, Cochrane and Fučeková, smartphones and MALL had an important role, but only from the viewpoint of the importance and the possible effectiveness of using smartphone applications in teaching and learning language skills in EFL context in the 21$^{st}$ century. They emphasized that smartphone applications need to be taken into consideration since they have become an inseparable part of learning process. The importance of using mobile phone applications for assessing speaking skills is relevant for my study. Unfortunately, studies on assessing speaking skills with smartphone applications were not found. Thus, my study will produce new information on the field of MALL and assessment.

Questionnaires as data gathering tools were used in these previous studies on the field of MALL. Fučeková (2018) used an electronic questionnaire while in Abugohar et al.'s (2019) and in Cochrane's (2015) exploratory study a mixed method was used. The topic of Abugohar, Yunus and Rashid's (2019) study is *Smartphone applications as a teaching and learning technique of speaking skill*. Their results were that teachers had high positive perceptions of using smartphone applications in teaching speaking skills and improving fluency, confidence and accuracy but that their classroom practices did not correlate with these high positive perceptions. Conversely, their classroom practices were insufficient, and they did not have enough experience of using smartphone applications. As a conclusion, Abugohar et al. (2019) stated that smartphone applications are efficient to improve fluency, confidence and accuracy in speaking skills and that teachers should be encouraged to use them efficiently and that intensive training should be provided to both teachers and students. The topic of Cochrane's (2015) study is *Activities and reflection for influencing beliefs about learning with smartphones*. The findings were that integrating new applications into English classes through tasks may be a way to introduce students how to use applications in a more productive way in learning and have a positive effect on their smartphone use from pleasure use to academic and learning purposes use. The topic of Fučeková's (2018) study is *Developing English skills by means of mobile applications*. In her study, the importance of smartphone applications as a learning method was demonstrated and the most used mobile device were smartphones. Her questionnaire contained questions about the type of mobile devices. In my questionnaire, I

asked about types of applications and how they could be used in assessing. In addition, in Fučeková's study, when using smartphone applications, learners seemed to be interested in speaking and pronunciation, even though listening was number one skill – which can also improve ones' speaking skills.

# 3 The present study

This chapter describes the design of the present study and the methods used in it, in a way that it can be replicated if necessary. First, I will describe my research problem and research questions. Secondly, I will tell about the participants and data collection. I will discuss the challenges there may be when selecting your sample and if all the selected participants do not complete the questionnaire, i.e. data reduction and subsequent return rate and their effect on validity and reliability. Moreover, I will describe the rationale behind the selection and collection of my data. Thirdly, I will discuss my method of data collection which was an online questionnaire and the advantages and disadvantages when using a questionnaire. As in the phase of choosing the sample, also the choices made regarding your questionnaire can cause bias and affect the results and the validity and reliability of the study. Furthermore, I will clarify the reasons why I chose the 21 questions to be the ones in my questionnaire. Finally, I will describe the methods used in the analysis and the differences and the similarities between analyzing closed ended and open-ended items.

## 3.1 Aim and research questions

The aim of my study was to investigate *how students perceive the use of smartphones in the formal assessment of speaking skills*. To do this, I chose to use an online questionnaire in Webropol, which is a platform for collecting and analyzing data. As Mackey and Gass (2005: 92-96) state, surveys are a common method to collect data on opinions, attitudes, perceptions and motivations, also when learners' opinions of themselves are of interest. The three research questions that my study sought answers for were: 1) *How would students like to be assessed formally in speaking skills?,* 2) *How would students like to be assessed formally in speaking skills when using smartphone applications?* and 3) *Which applications that they already know could be used in assessing speaking skills in their opinion?*

**3.2 Participants and data collection**

The data was collected in an upper secondary school in Eastern Finland after I had asked permission to carry out my study from the principal and the English teachers. First, I sent the data protection notification to the principal in November 2019. After having received the permission, I sent a link and a qr-code of my online Webropol questionnaire to the teachers and asked if they could have the students answer it during their English lessons, which they did in December 2019.

Overall, the questionnaire was opened 54 times, started 41 times and finished 33 times, i.e. eventually there were 33 questionnaires in Webropol for me to be analyzed. Dörnyei and Taguchi (2010: 63-64) mention the problem of participant self-selection which could also be seen in my study. We can see here the effect of the participants' own decision not to complete the questionnaire even if they were selected systematically to be part of this study. Since the questionnaire could not be compulsory and due to the data protection notification, the respondents were allowed choose whether they participated or not and whether they finished the questionnaire or not. The validity of my study would have been better if all the 54 participants who opened the questionnaire would also have finished it. Unfortunately, this was not the case. Respondent motivation and subsequent return rate can affect the validity of surveys, as Dörnyei and Taguchi mention (2010: 63-64). When purely statistical point of view is considered, the sample should be 30 or more participants, but smaller samples can be compensated by using non-parametric procedures (Hatch and Lazaraton 1991, as quoted in Dörnyei and Tagushi 2010: 62). But since certain statistical procedures require more than 50 people, or a minimum of 100 participants, the present study does not have statistical significance (Dörnyei and Tagushi 2010: 62-63).

In fact, Tolmie, Muijs and McAteer (2011: 18-22) also discuss the process of sampling and its effect on data reduction. They state that because of selection processes during a study, it is not possible to collect data without data reduction or simplification. Sampling is one of these phases of selection during the process of carrying out a study. Another important phase is data analysis. Therefore, both sampling and data analysis affect the results. Tolmie et al. (2011: 18-22, 287-303) and Dörnyei and Taguchi (2010: 63-64) underline that the participants may be more motivated than those who did not choose to participate, and this can affect the results and the fact that exactly these respondents are representative of the wider population and not some others. The results of a study are based on the responses of the participants and the analysis

made by the researcher. Sampling bias means that in the results there can be over- or under-representation of cases.

When I was selecting my sample and participants, upper secondary school students were chosen since they probably have more experience on the assessment of speaking skills in English and since I supposed that I would have more reliable and valid answers with them than with younger pupils, e.g. in an upper comprehensive school. I did not use a random sample of students since I preferred the groups to be third-year students. The aim of the selection was to have more reliable and valid information from third year students who have already studied more English in the upper secondary school level. But since not all the sampled third-year students finished the questionnaire, there were also second-year students as respondents. Thus, a convenience or opportunity sampling was used, since the main reason for selection was the convenience for the researcher (Dörnyei and Tagushi, 2010: 60-62). In other words, participants were selected because of easy accessibility and availability at a certain time. Moreover, the sampling was purposeful, since the participants were of specific age which was a key characteristic for the purpose of my study (Dörnyei and Tagushi, 2010: 60-62).

## 3.3 Questionnaire

There are many advantages with questionnaires. As Mackey and Gass (2015:92-96) state, the advantages of a questionnaire are that it is more practical and economical than individual interviews. Moreover, they state that the number of respondents can be bigger and data gathering process can be made in many flexible ways. They also suggest using oral answers which could be recorded and be a good choice for those who have limited literacy, but in this study, oral answers were not used.

However, there are also problems with questionnaires. Mackey and Gass (2005: 92-96) discuss inaccurate or incomplete responses and bias. I will first discuss inaccurate or incomplete responses and the reason to use L1 in my questionnaire and then bias. There were 21 questions in the questionnaire in Finnish. The reason why I chose Finnish and not English was that Mackey and Gass (2005: 92-96) state that one problem which is related to the analysis of questionnaire data are inaccurate or incomplete responses. They state that this problem occurs when using L2 in questionnaires, especially in open-ended items if they are not understood or if respondents feel uncomfortable when writing the answers. Therefore, the students' native language was a better choice in my questionnaire than L2 would have been. Even though there

was one female respondent who had another L1 than Finnish, she did not mention what her mother tongue is and whether it affected her responses is therefore difficult to know.

To avoid bias and maximize the efficiency of my questionnaire, my purpose was to have simple formats and unambiguous questions as Mackey and Gass (2015:92-96) propose. Moreover, they state that the format should be user-friendly and the questions clear. In Webropol, it is possible to review the format of the questionnaire on a computer, tablet and smartphone. Furthermore, Mackey and Gass propose that questionnaires should be reviewed and piloted to avoid bias. The pilot version of my questionnaire was reviewed by my supervisor and my MA thesis group, and the final version was tested in Webropol by two upper secondary school teachers and one student from a university of applied sciences. These versions were modified according to the comments received in the revision and the testing phases. Conversely, it was not piloted among the research population, in this case among students on upper secondary schools, which would have been advisable according to Mackey and Gass (2015:92-96).

Now, I will describe the questionnaire in detail (see Appendix). On the first page there was a data protection notification and a description of the study and its aim. In the section of background information questions, the only personal information items asked from the respondents were the year of studies, sex/gender, native language and the latest grade in English in upper secondary school (questions 1-4). Thus, the only demographic information items asked were sex/gender and L1.

There were both closed ended and open-ended items in the questionnaire. Closed ended items offer quantitative data and open-ended items produce mostly qualitative data. Regard to reliability, a closed-ended question is more reliable since an open-ended question can be answered in any manner when respondents want to express their own ideas and thoughts and there can be unexpected data to be analyzed (Mackey and Gass, 2005:92-96). As Dörnyei and Taguchi (2010: 83-110) state, closed-ended questions are most common in questionnaires. In my questionnaire there were fifteen closed ended items (questions 1-7, 10-12 and 17-21) and only six open-ended items (questions 8-9 and 13-16).

Dörnyei and Taguchi (2010: 26-38) divide closed-ended items into five categories: rating scales, multiple-choice items, rank order items, numeric items and checklists. The closed-ended items in my questionnaire were all multiple-choice items, two of which were kind of rating scales. These rating scales were more likely semantic differential scales, since in them it was asked about tension experienced in exams in general or in exams in speaking skills on a scale from 0 to 10 in which 0 is *not at all* and 10 is *really a lot* (questions 17 and 18).

Open-ended questions can be divided into four categories: specific open questions, clarification questions, sentence completion items and short-answer questions (Dörnyei and Taguchi 2010: 26-38). In my questionnaire, the open-ended items can be defined as specific open questions or short-answer questions and the open-ended items appearing in closed ended items were all specific clarification questions. I will here use Dörnyei and Taguchi's (2010: 26-38) definitions of the used three categories of open-ended questions. Specific open questions ask about concrete pieces of information and the answers can be followed by a "Why?" question. Answers in short-answer questions are usually more than a phrase and less than a paragraph if they are well formed. They deal with only one concept or idea. The purpose of clarification questions is to have more clarification, e.g. in a form of "please specify".

In my questionnaire questions 8-9 and 13-16 could be categorized both in the groups of specific open questions and short-answer questions. Conversely, in the questionnaire, the clarification questions were used in some of my closed-ended items in the following way: in question 3 if L1 was other than Finnish; in questions 5-7 if the answer was *yes* when asked about previous speaking skills exams; in questions 10-12 if students proposed (also) other smartphone applications to be used in speaking skills exams than the options in ready-made suggestions or if they did not want to use smartphone applications at all in speaking skills exams; in question 19 a clarifying why-question if they thought that using smartphone applications would reduce or increase tension in speaking skills exams; and in questions 20-21 also a why-question, if they thought that students or teachers would or would not need training for knowing how to use smartphone applications in assessing situations.

The previous studies, as discussed in chapter 2.3 about MALL, gave me some ideas to formulate my own questionnaire. Firstly, the last two questions, 20 and 21, were based on Abugohar et al.'s (2019) findings that more technical support and training – both for teachers and students – were needed and that support and training were considered as an important issue to increase motivation to use smartphone applications. Secondly, the applications used in Abugohar et al.'s (2019) study include real life, i.e. authentic, target language use (TLU) possibilities. In my questionnaire, the purpose of questions 10-15 was among other things to reveal TLU applications. Moreover, the categories of smartphone applications in questions 10-13 were slightly based on Abugohar et al.'s way to divide smartphone applications into three categories: 1) speech-to-text transcription applications 2) audio recording animation-based applications and 3) automatic speech analysis video-based applications, and on the task type of voice recorders in Cochrane's study (2015). Thirdly, Cochrane used reflection in homework

tasks as one component and in my questionnaire, I used SWOT analysis as a reflection tool in questions 13 (strengths and weaknesses) and 15-16 (opportunities and threats).

Some of the open items of my questionnaire were my research questions (RQs). RQ1 was asked in two questions, in the obligatory question 8 *If you could decide, what kind of speaking skills exams there would be?* and in the voluntary question 9 *How would you liked to be assessed/evaluated in speaking skills exams in English?* RQ2 was asked in the obligatory closed items 10-12 which were questions about recording-, drama- and role/avatar-applications. RQ3 was asked in the obligatory open-ended item 14 which was about applications students already use and which ones of those they think could be used in the assessment of speaking skills. Moreover, in the obligatory closed items 10-12, respondents were asked if they knew some other recording-, drama- and role/avatar- applications than those already mentioned.

## 3.4 Methods of analysis

This study was empirical and mixed methods were used since both quantitative and qualitative methods were employed. A data-driven analysis was used since the data was analysed inductively into categories.

Dörnyei and Taguchi (2010: 83-110) state that when processing and coding the data, each questionnaire should be marked with an identification number. Conversely, in my study the questionnaires in Webropol were not marked individually with any identification numbers since I used a public link and not e-mail or text message to participate in the survey. The participants were able to open the questionnaire using a qr-code their teachers gave to them. Therefore, the students could be identified in a way individually only by the time on which they sent their questionnaire in Webropol and this time could be same for several students.

Closed ended items can be easily quantified and analyzed (Mackey and Gass, 2005:92-96). Since I was using Webropol, the data of closed ended questions was quantified by the program. Numerical questionnaire data can be processed by statistical procedures and coding phase (Dörnyei and Taguchi, 2010: 83-110). Thus, answers were converted into numerical scores, e.g. gender data, by Webropol. As the types of instrumentation, I present figures and tables and some verbal descriptions of the data. Moreover, there are quantitative data discussions in which I compare the answers between men and women (Dörnyei and Taguchi, 2010: 83-110).

Open-ended questions were processed by the methods of systematic qualitative content analysis which produced both quantitative and qualitative data based on which categories were done, summaries were written, and tables and figures were compiled (Dörnyei and Tagushi 2010: 83-110). According to the types of data, parametric and non-parametric procedures were used (Dörnyei and Tagushi 2010: 83-110). Parameters are used when you want to describe numerically some characteristics of your participants (Mackey and Gass 2005: 362). In addition, examples of the relevant answers in open-ended questions are usually provided in the results section.

As Dörnyei and Tagushi (2010: 83-110) suggest, in specific open questions, I have a limited number of categories. I made decisions how to label two similar responses into one category, for example in question 8 which is based on the RQ1 *How would students like to be assessed formally in speaking skills?,* the open answers *a speech (puhe)* and *a presentation (esitelmä)* were labelled into the same category even if they were not completely identical responses. Hence, I made qualitative interpretation having the research problem and questions in mind. As Dörnyei and Tagushi (2010: 83-110) underline, it is important not to overgeneralize, but it is obligatory to generalize. They state that if parametric procedures are used too much, there may be overgeneralization. In clarification and short-answer questions there are more subjective elements. Thus, I used content analysis and reduced data in a reliable manner.

Since my open-ended questions were relatively clear to answer to, it was quite easy to label the answers into categories. The most problematic issue was the answers in which it was easy to see that the respondents had not understood the question, had not answered to the question, or had answered in an inappropriate way. For example, when analysing question 7 in which it was asked if the students had had formal exams on speaking skills in upper secondary school and what kind of exams they had had if they answered *yes*, there were a few unclear answers and therefore I asked clarification by e-mail from one of their English teachers about the exam they have on their ENA8 course which is a course on speaking skills. All in all, unclassified elements are reported as such or they are not taken account (Dörnyei and Tagushi 2010: 83-110). In closed ended items 10-12 and 19, I compared the answers according to sex/gender (question 2). Conversely, I did not take into consideration the effect of other background information (questions 1, 3-4): the year of studies, the latest grade in English in upper secondary school or native language. The reasons for this decision were the following: 1) having both 2$^{nd}$ year students (30%) and 3$^{rd}$ year students or older (70%) was not my own

choice but obligatory to have enough participants, 2) only three respondents (9%) had grades between 5-6 and 3) there was only one student who did not have Finnish as her L1.

## 4 Analysis of results

In chapter 4.1, I will present the respondents and their previous exams on speaking skills. Chapter 4.2. will then present the suggestions respondents made about different kind of exam types and smartphone applications. Chapter 4.3. will introduce the SWOT analysis on smartphone applications when used in speaking skills exams, as experienced by the respondents. The SWOT analysis will also focus on tension in exams and the effect of smartphones on it as well as the need for training when using smartphone applications in speaking skills exams. Chapters 4.2 and 4.3 will also reveal some differences in the results through the background variable of gender/sex.

**4.1 The respondents and their previous exams on speaking skills**

Altogether, there were 33 respondents, out of whom 23 were women (70%) and 10 men (30%). One of the 33 respondents had another native language than Finnish, but she did not mention which it was. The proficiency level of the respondents was mostly good (grades 7-8) or excellent (grades 9-10). Only three respondents (9%) had grades between 5-6 in their last English course in upper secondary school, while 15 (45,5%) had grades between 7-8 and another 15 (45,5%) had grades between 9-10.

When asked if the respondents had had formal exams on speaking skills in lower and upper comprehensive school and in upper secondary school (questions 5-7), the answer was 'yes' only in 24% of the answers concerning lower comprehensive school and in 46% concerning upper comprehensive school. The result of 'yes' answers was even higher when asked about exams in upper secondary school: it was 76%. Therefore, the respondents have had more exams on speaking skills on upper levels than in lower levels. But when taking also into consideration the answers 'no' and 'I do not remember/I cannot say', this difference might be explained with the fact that the respondents do not remember as clearly their lower comprehensive school experiences than those in their current school – the upper secondary school. Namely, when asked about exams in lower comprehensive school, 40% answered clearly 'no' and 36% answered 'I do not remember/I cannot say'. When asked about exams in upper comprehensive

school, the result with 'no' answers was 39% of the respondents, and there were clearly less those who answered 'I do not remember/I cannot say', only 15%. In upper secondary school, only 15% answered 'no' and 9% 'I do not remember/I cannot say'.

If the respondents answered 'yes' when asked about previous exams on speaking skills (questions 5-7), they had an opportunity to answer to a clarification question 'what kind of exams they were'. All the answers for this clarification question are summarized in Table 1. When asked about exams in lower comprehensive school all but one of the eight respondents had answered that they *read aloud a text* (83%). The one who did not have the same answer had answered *a discussion* (17%) and more precisely *a discussion with his/her teacher*. In upper comprehensive school there were already more methods in speaking skills' exams even though the options *read aloud a text* (15%) and *a discussion* (21%) were still common methods. In addition to these methods there was *a presentation or a speech* (50%), which was the most common method in upper comprehensive school. Moreover, there was one who had answered *recording (nauhoittaminen)* and one who had answered *paper exams* which was an unclear answer. When moving to upper secondary school, the methods used started to differ and multiply even though the most common method was *a presentation or a speech* (53%) as it was also in upper comprehensive school. In upper secondary school they have the course ENA8 on speaking skills and the second common answer was therefore *the final exam of the course ENA8* (23%). One of the respondents had clarified that in the exam there are different parts which are *reading aloud, translating a text in your own words* and *a discussion*. Moreover, I sent an e-mail to one of their English teachers to ask for explanations for the answers of this clarification question 'what kind of exams they were' (question 7) and this teacher explained that the exam is 20 minutes long and teachers choose pairs being quite on a same proficiency level. The exam is recorded or filmed by teachers for assessing and evaluating purposes. The third common answer was *recording (nauhoittaminen) or filming (videointi)* (8%) and there can be three different types of situations in which this may happen. First, *recording (nauhoittaminen) or filming (videointi)* can be a part of the final exam of the course ENA8. Secondly, the English teachers told in the e-mail that students who are too afraid of being in front of the class for presentations or speeches can have permission to tape their presentations only for teachers to see them. Thirdly, they have some exams of this kind. Furthermore, one respondent answered that they have *oral exercises in pairs which their teacher evaluates*. In addition, there were irrelevant answers or answers which show that the respondent had not understood the question (12%) and these questions were one reason for my clarification e-mail for their teacher. All in

all, there were only two respondents who did not use the opportunity to answer to the clarification question 'what kind of exams they were'.

Table 1. Respondents' previous exams in speaking skills.

| Previous exams in speaking skills | Lower comprehensive school N=6 | Upper comprehensive school N=13 | Upper secondary school N=23 of which one in two categories |
|---|---|---|---|
| *Read aloud a text* | 5 (83%) | 2 (15%) | |
| *A discussion (with his/her teacher)* | 1 (17%) | 3 (21%) | |
| *A presentation or a speech* | | 7 (50%) | 14 (53%) |
| *Recording* | | 1 (7%) | |
| *The final exam of the course ENA8 (parts: reading aloud, translating a text in your own words, a discussion)* | | | 6 (23%) |
| *Recording or filming* <br> *1) a part of the final exam* <br> *2) a form for those who are too afraid of being in front of the class* <br> *3) actual recording or filming exams* | | | 2 (8%) |
| *Oral exams in pairs which their teacher evaluates* | | | 1 (4%) |

## 4.2 Formal assessment of speaking skills

4.2.1 Respondents' suggestions of different types of exams

Questions 8 and 9 were concerned with the formal assessment of speaking skills. Question 8 was an open-ended item in which the respondents could tell what kind of exams they would like to have. The answers were categorized in two ways. The first categorization was made by the task type and the second categorization was made according to the implementation, i.e. if they preferred have exams alone with the teacher, with (an)other student(s) or in front of the class or not.

As there were more suggestions for the different kinds of exams (N=40) than respondents (N=33), the percentages were calculated by dividing the number of the answers (see Table 2). The most wanted task type was *discussion* (28%). The two task types which shared the second place were *reading aloud* (15%) and *a speech/a presentation* (15%). The

third task type was *a recording* (8%). Moreover, there were a few who wanted to have *speaking* tasks (5%) or tasks of *asking questions and answering or only answering questions* (5%). In addition, there were answers which were not categorizable to any task type, e.g. *I wouldn't like to have exams in which you have to be in front of the class, I prefer those with your teacher, just the two of us (En haluaisi luokan edessä pidettäviä, mieluummin opettajan kanssa keskenään.).* Furthermore, there were answers which had a really general or universal aspect, e.g. *basic stuff (perusasioita);* answers which were satirical, e.g. *kind of exams which do not exist (sellaisia, joita ei olisi ollenkaan);* and answers which showed that the respondent had not understood the question, e.g. *eight-nine (kasi-ysi).*

Table 2. Respondents' proposed exam categories according to the task type and implementation (question 8).

| 1) Task type | N=40 |
| --- | --- |
| *A discussion* | 11 (28%) |
| *Reading aloud* | 6 (15%) |
| *A speech/a presentation* | 6 (15%) |
| *A recording* | 3 (8%) |
| *Speaking* | 2 (5%) |
| *Asking questions and answering or only answering questions* | 2 (5%) |
| | |
| **2) Implementation** | **N=41** |
| *In pairs* | 9 (22%) |
| *With their teacher* | 4 (9%) |
| *In front of the class* | 5 (12%) |
| *Not in front of the class* | 2 (5%) |

The percentages in the second categorization of implementation were again calculated by dividing the number of the answers (N=41) for this question (see Table 2). The most wanted way to have an exam was *in pairs* (22%). There were also several who wanted to have exams *with their teacher* (9%). Additionally, those who had mentioned *reading aloud* (7%) could be added to this same group since it can be supposed that *reading aloud* would happen with your teacher listening to you. In addition, there were some other answers (19%) which could not be clearly categorized since the manner was not clearly expressed. For example, *discussion, asking questions and answering or only answering questions* and *speaking exams* could be held in pairs, in groups or with the teacher. The task type of *a recording* (here 7%) could be done alone, in pairs and in groups or with the teacher. The most interesting division was between those who wanted to be *in front of the class* (12%) and those who did *not* want to do so (5%). Those who would like to be *in front of the class* mentioned task types as *a speech/a presentation*. Those

who did *not* want to be in front of the class mentioned it clearly, e.g. *not in front of the whole class (ei koko luokan edessä)* and *I would not like to have exams in front of the class, I prefer exams with the teacher alone (En haluaisi luokan edessä pidettäviä, mieluummin opettajan kanssa keskenään.)*.

As Figure 1 shows, when the respondents were asked how they would like to be evaluated in speaking skills exams in English, 44% of the respondents wanted to have a verbal evaluation, while 34% wanted to have both a grade and an explanation of their proficiency. 16% wanted only to have a grade. One respondent wanted to have at least a verbal evaluation and for one respondent everything goes.



- a verbal evaluation 44%
- both a grade and an explanation of their proficiency 34%
- a grade 16%
- at least a verbal evaluation 3%
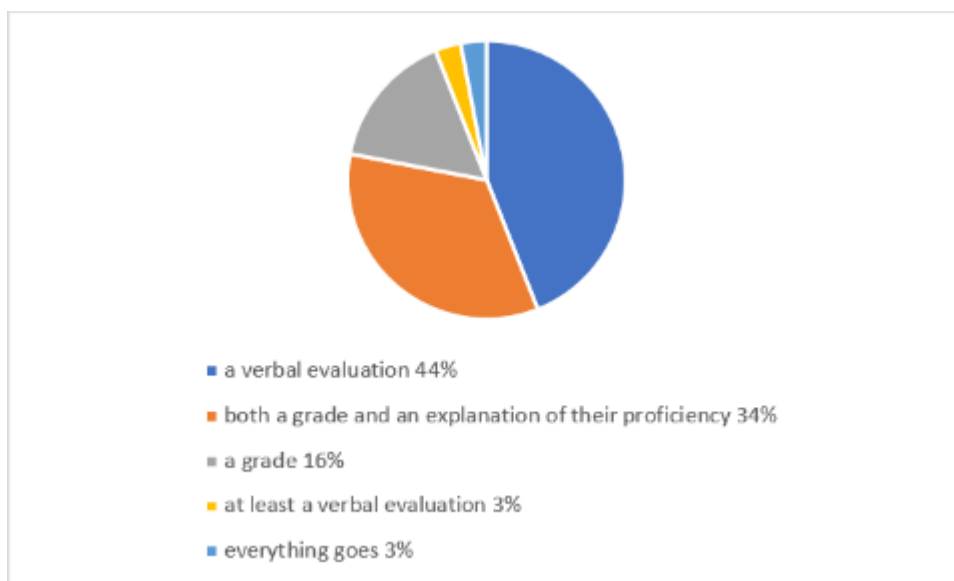- everything goes 3%

Figure 1. The way respondents would like to be evaluated in speaking skills exams in English (question 9). N=32.

The respondents who wanted to have a verbal evaluation had for example the following answers: *I would like to know what should be learned more and what is already fine (Haluaisin tietää, mikä kaipaa kehittämistä ja mikä menee hyvin.)* and *The teacher could tell me how it went and what should be improved (Öö no vaikka kertoo vaan että miten meni mitä pittää parantaa).* In other words, it is important for them to know how to improve their speaking skills and that teachers use assessment *for* learning, not only assessment *of* learning (Keurulainen 2013: 37-45).

4.2.2 Respondents' suggestions of the use of smartphone applications

Questions 10-12 and 14 focused on smartphone applications and their use in the assessment of speaking skills in English. In the closed ended questions 10-12 the respondents had an opportunity to choose more than one of the given options for applications and to answer to two open-ended items: 'Which one(s)?' when asked about applications and 'Why not?' if they did not want to use them.

In question 10 the students had an opportunity to tell which voice recording applications of their mobile phone they would like to use in exams on speaking skills in English. 76% of the respondents chose 'Recording applications for speech' and 45% chose 'Whatsapp'. Women chose more often these two options than men did: 83% of women and 60% of men chose 'Recording applications for speech' and 57% of women and 20% of men chose 'Whatsapp'.
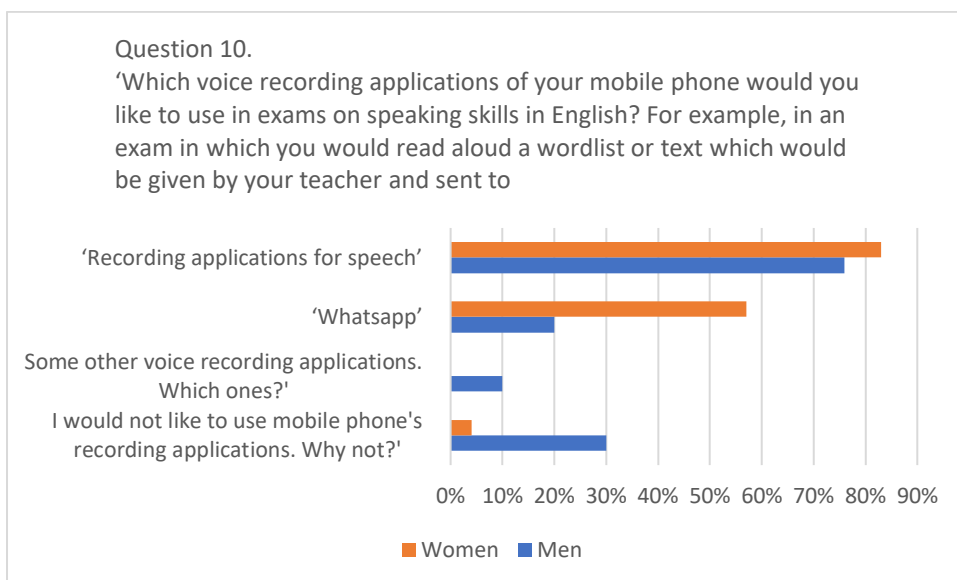


Figure 2. Question 10. Answers compared between women and men.

In question 10, 12% of the respondents would not like to use recording applications. The difference between women's 4% and men's 30% is clear. There were four responses to the open-ended item 'Why not?'. Men's responses were the following: *You could try several times until you get a good one and then your direct knowledge/skills on oral skills wouldn't be proved. (Voisi yrittää monesti uudelleen kunnes saa hyvän jolloin suora osaaminen ei näy.); Face-to-face works better. (Faxe to face toimii paremmin. –* respondent's typo*) and It's not the mobile phone that will help you to learn at school. (Ei se kännykkä siihen koulunkäyntiin auta).* The one answer from a woman was *I don't like oral stuff. (En pidä suullisista jutuista.)*
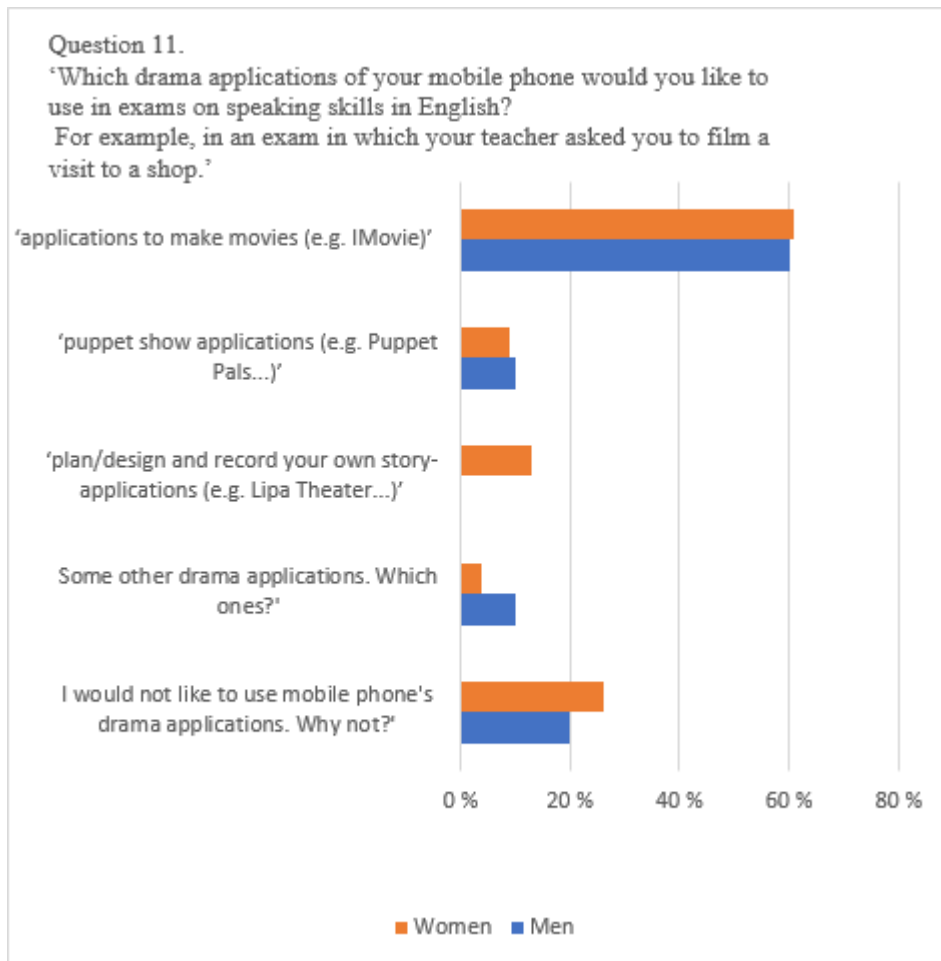
Figure 3. Question 11. Answers compared between women and men.

In question 11 the students had an opportunity to tell which drama applications of their mobile phone they would like to use in exams on speaking skills in English. The most popular option was 'applications to make movies (e.g. IMovie)' (61%). The following two options got both 9% of answers: 'puppet show applications (e.g. Puppet Pals...)' and 'plan/design and record your own story -applications (e.g. Lipa Theater...)'. When comparing women and men (see Figure 3), none of the men chose the option 'plan/design and record your own story -applications (e.g. Lipa Theater...)' while 13% of women did choose it. But 10% of men would be willing to try 'some other drama applications' while only 4% of women would like to try other drama applications. There were more women (26%) who would not like to use mobile phone's drama applications than men (20%). Conversely there were no significance difference between women (61%) and men (60%) in choosing 'applications to make movies (e.g. iMovie)' or 'puppet show applications (e.g. Puppet Pals...)', (women 9% and men 10%). It could be concluded that both women and men are equally ready to make movies and puppet shows, while only women want to make their own stories. More women would not like to use applications at

all while men are more open to try also other drama applications. The reasons for these differences could be that men are more self-confident when using new technology and applications and that they are more open to new opportunities. While women might be more open to use their imagination and tell stories. But men also are ready to make puppet shows and thus ready to use their imagination.

In question 12 the students had an opportunity to tell which role or avatar applications of their mobile phone they would like to use in exams on speaking skills in English. 39% of the respondents would like to use 'Avatar applications (e.g. Morfo)'. 12% would like to use 'plan/design and record your own story (e.g. Lipa Theater...)' applications and 9% would like to use 'puppet show applications (e.g. Puppet Pals...)'. Women (43%) would like to use 'Avatar applications (e.g. Morfo)' more often than men (30%) and men would not like to use the two other given options at all. But 20% of men would like to use 'some other drama applications', the option which none of the women chose. As it was already the case in question 11, men seem to be more open to some other/new drama applications while women are more interested in puppet show and storytelling applications.
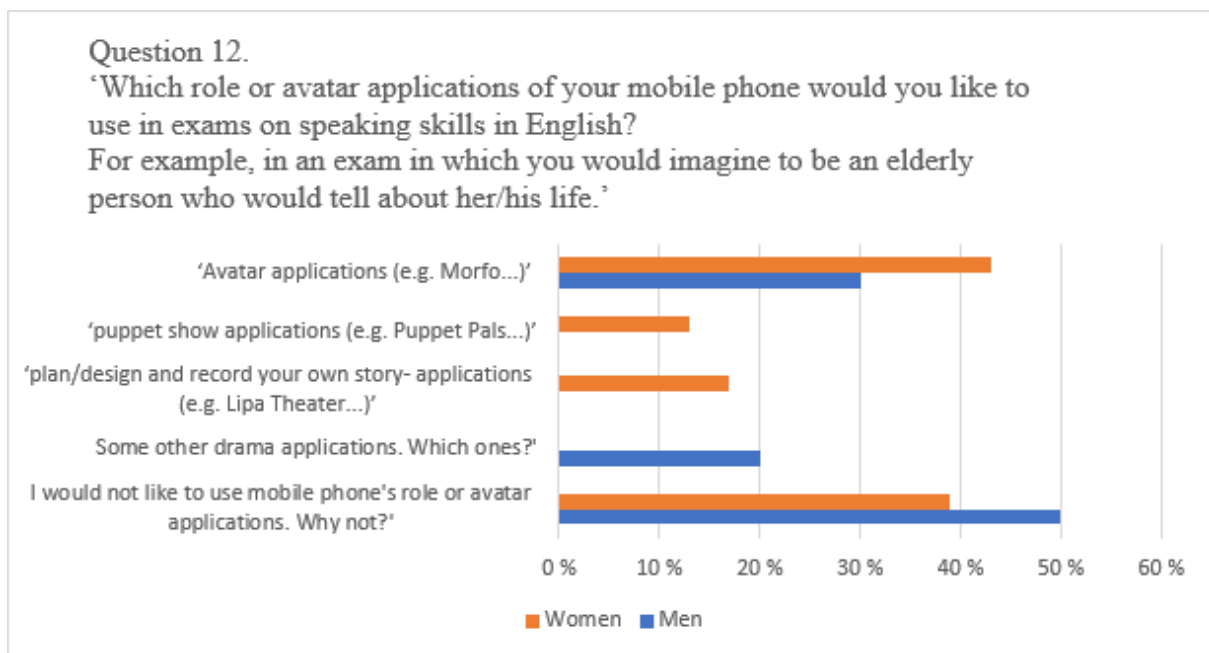


Figure 4. Question 12. Answers compared between women and men.

Almost a quarter of the respondents in question 11 would not like to use drama applications (24%) and almost half of the respondents in the question 12 would not like to use any role/avatar applications at all (42%). The reasons why they would not like to use drama/role/avatar

applications were divided in two main categories: 1) students' experiences of these applications and 2) the properties of these applications. We will first see the answers for the first main category. 20% of the respondents think that they are difficult to use and/or it takes time to use them. 13% feel that it is somehow uncomfortable to use them. For example, *Filming myself doesn't feel comfortable when someone is going to evaluate it. (Itseni kuvaaminen ei tunnu mukavalta, kun joku arvioi sen.).* 13% of the respondents also doubted the usefulness of the applications. For example, *I don't think they are useful. (Ei niistä taida olla hyötyä.).* In the second main category, the properties of applications, one respondent doubted if they would work while another respondent was optimistic and thought that she would probably animate them herself. Moreover, 26% of the respondents told that they have not used these applications at all. For example, *what are they (mitä ne ovat)* and *I can't, and I don't even know what they are (en osaa, enkä tiedä edes mitä ne on.)*

A list of the mobile phone applications the respondents proposed to be used in the formal assessment of speaking skills are presented in Table 3. In question 10 only one respondent suggested other applications for voice recording than those already given in closed ended items. None of the women proposed anything. In both questions 11 and 12, two of the respondents proposed some other drama applications and role/avatar applications. The other options proposed in open-ended items were the following: *Snapchat* for voice recording, smartphone's own *camera* and *Gacha life* for making drama, and *Gacha life* and *Bittmoji* as role/avatar applications. But after having tried *Gacha life*, in my opinion, it can only be used in written form.

In question 14, the respondents could tell which of the mobile phone applications that they had already used could be used in exams on speaking skills (see Table 3). There were 39 different kind of answers. The applications were categorized as those which are only voice recording applications (13%) and those applications which incorporate both voice and picture/images (49%). In the answers the applications with which only voice could be recorded were *dictation/tape recorder (sanelin/ääninauhuri)* and once *Whatsapp, to record a voice message to the teacher (Whatsapp, äänittää opettajalle ääniviestillä).* The applications having both voice and picture/images were *Whatsapp* (23%) (in these cases the answer *Whatsapp* was interpreted as both), *Snapshat* (10%) and *mobile phone's video camera or camera* (8%) (*camera* was interpreted as *video camera* since there are no sounds/voice in a normal camera). The following applications which incorporate both voice and picture/images were mentioned only once: *Gacha life* and *in Gacha life performance/presentation and dubbed dialogues*

*(Gacha lifessa esitys ja päälle puhuttut dialogit); applications for processing videos (videonmuokkaussovellukset)* without no names mentioned and *Twitter.*

As we can see in the table 3, among the suggested mobile phone applications (questions 10-12, 14) there are applications which could be used in the formal assessment of speaking skills and others which could not be used for this purpose since there is no function for oral skills to be presented when only text and/or pictures are used.

Table 3. A list of the mobile phone applications the respondents proposed to be used in the formal assessment of speaking skills (questions 10-12, 14).

| MOBILE PHONE APPLICATIONS which could be used in the formal assessment of speaking skills | MOBILE PHONE APPLICATIONS which could <u>not</u> be used in the formal assessment of speaking skills |
|---|---|
| **VOICE**<br><br>• *dictation/tape recorder* | **TEXT AND PICTURE**<br><br>• *Gacha life: only in a written form as a presentation; you can fill in the speech bubbles speaking but the teacher could not hear that part of the process* |
| **VOICE AND PICTURE**<br><br>• *Snapchat, Whatsapp, Twitter*<br>• *applications for processing videos*<br>• *camera, if it is supposed to be interpreted as a video camera* | **PICTURE**<br><br>• *Bittmoji* |

In 36% of the answers the respondents did not know any applications to be used in exams on speaking skills, did not know how to use them, did not want to use them or did not 'do' exams on speaking skills. There was one respondent who did not understand the question since he/she responded: *Probably actually, a vocabulary application e.g. Otava vocabularies...* (*Varmaankin oikeastaan, jokin sanastosovellus esim. Otava sanastot...*).

## 4.3 Students' perceptions of the use of smartphone applications

4.3.1 Strengths, weaknesses, opportunities and threats (SWOT) of smartphone applications

Question 13 asked about strengths and weaknesses of the (voice) recording, drama, role and avatar applications that the respondents had used, and thus were able to evaluate and judge their

applicability in the assessment of speaking skills. As it can be seen in Table 4, the respondents mentioned more often weaknesses (20%) than strengths (16%).

There were 25 different kind of answers to the question 13 and the percentages were counted using the number of answers. As weaknesses the respondents mentioned 'the difficulty to use these applications' (8%) and 'Movie making applications' were mentioned separately as difficult to use (4%). 8% of the respondents were worried about 'the quality of the voice' when using mobile phones. One of the respondents mentioned as a weakness that ...*when using voice messages in whatsapp they may be sent too early when touching the recording button (Koulutöissä ei ole vielä tullut näitä käytettyä, mutta esimerkiksi whatsapissa äänitettäessä ääniviesti lähtee vahingossa liian aikaisin, kun liikauttaa äänityspainikkeelta.)* As strengths the respondents mentioned 'creativity' (8%) and 'easiness to use these applications' (8%). An example of 'creativity' is the following answer: *You can be free and create your own personality/person (Niissä voi olla vapaa ja luoda persoonaa).* There was a common axis 'usability' between the strengths and weaknesses, and it was interesting that there were as many respondents who found these applications difficult to use as those who found them easy to use. An example of 'easiness' is the following answer: *Usually, recording applications are handy to use (Äänittämis-sovellukset yleensä käteviä).*

Table 4. SWOT: strengths and weaknesses of using smartphone applications in the assessment of speaking skills (question 13).

| **Strengths** | **4 (16%)** |
|---|---|
| *creativity* | *2 (8%)* |
| *easiness to use these applications* | *2 (8%)* |
| **Weaknessess** | **5 (20%)** |
| *the difficulty to use these applications* | *2 (8%)* |
| *the quality of the voice* | *2 (8%)* |
| **Strengths/ Weaknessess** | **2 (8%)** |
| *"You can record several times and in parts".* | *1 (4%)* |
| *"They are not face-to-face situations".* | *1 (4%)* |

8% of the answers were ambivalent and could not be categorized as a strength or as a weakness. This answer was a same kind of answer as in questions 10-12 as an answer to the open-ended item 'Why not?' want to use applications in exams: *You can record several times and in parts*

*(voi nauhoittaa monta kertaa ja osissa)* and thus it could be seen as a weakness but the use of the modal auxiliary *can (voi)* makes the categorizing of it difficult. Moreover, the following answer can be categorized in both categories: *they are not face-to-face situations (ne ei oo kasvotusten).* It depends how the respondent feels about face-to-face situations. It would have been interesting to conduct interviews with the respondents to have deeper knowledge about these ambivalent responses.

52% of the respondents did not know at all or they knew hardly these applications or had not used these applications at all. This is interesting since it could be supposed that young people use their mobile phones a lot and thus, they would know various applications of these kind. One of these respondents mentioned that *I haven't used any of these except voice recording applications. (En ole käyttänyt mitään näistä paitsi äänittämis-sovelluksia).*

Table 5. Respondents' opinions of the new **opportunities** that using mobile phone applications in the formal assessment of speaking skills in English could provide (question 15, N=26).

| 1) The significance/importance of the applications for students | |
| --- | --- |
| a) the tension would vanish or diminish | 5 (19%) |
| b) the easiness of making, processing and transmission/sending | 5 (19%) |
| c) diversity/variety | 2 (8%) |
| d) creativity | 2 (8%) |
| e) easiness to be yourself | 1 (4%) |
| 2) The significance/importance of the applications for teachers | |
| Possible to repeat the performance/fulfilment of students and thus the evaluation/assessment will be easier, more accurate and more diverse/versatile. | 5 (19%) |
| Answers not categorized in 1) and 2) | |
| New possibilities | 4 (15%) |
| Unclear and empty answers | 2 (8%) |

Question 15 asked about what new opportunities students (N=22) thought there could be when using mobile phone applications in the assessment of speaking skills in English (see Table 5). The percentages were once again calculated by dividing the number of the answers (N=26) for this question. 77% of the answers were categorized into two main categories: 1) the significance/importance of the applications for students (58%) and 2) for teachers (19%). Conversely, 23% of the answers were miscellaneous and thus they were not divided into the main categories. The two main categories have subcategories. When considering 1) the opportunities for students, there were five subcategories: a) the tension would vanish or diminish (19%); b) the easiness of making, processing and transmission/sending (19%); c) diversity/variety (8%); d) creativity (8%) and e) easiness to be yourself (4%=one answer). The

significance these applications have for 2) teachers is according to the responses (19%) that it is possible to repeat the performance/fulfilment of students and thus the evaluation/assessment will be easier, more accurate and more diverse/versatile. 15% of the respondents mentioned 'new possibilities' which can be interpreted for both teachers and students or only for teachers or students. In addition, there were unclear or 'empty' answers (8%).

In question 16, it was asked about what kind of threats students thought there could be when using mobile phone applications in the assessment of speaking skills in English (Table 6). There were 24 different kind of answers which were divided into two main categories: 1) technical know-how and problems (17%) and 2) psychological problems (29%). 17% of the respondents did not have any ideas of possible threats and the rest of the answers (37%) were isolated or unconnected answers as shown in Table 6.

Table 6. Respondents' opinions on **threats** there could be when using mobile phone applications in the formal assessment of speaking skills in English (question 16, N=24).

| | |
|---|---|
| **1) technical know-how and problems** | 4 (17%) |
| **2) psychological problems**<br>a) the possibility to cheat<br>b) the problems of concentration and focusing | 7 (29%) |
| **Other options** | |
| Not having any ideas of possible threats. | 4 (17%) |
| Isolated or unconnected answers. | 6 (37%) |

The examples in 1) technical know-how and technical problems were *not knowing how to use can hinder using them (huono tekninen osaaminen voi haitata)* and *technical problems (tekniset ongelmat)*. The subcategories in 2) psychological problems were a) the possibility to cheat (21%) and b) the problems of concentration and focusing (8%). For example, in a) *maybe not taking so seriously (Ei oteta ehkä niin vakavasti)* and in b) *you may pay attention to something not relevant, such as to your mobile phone (ne saattavat viedä huomion johonkin epäolennaiseen, kuten siihen kännykkään).* The isolated answers mentioned problems in interaction, the absence of TLU and authenticity, the bad quality of the voice and its effect on evaluation (teacher aspect), and mobile phones not being the best choice (pessimistic ones).

4.3.2 The effect on tension in formal exams

Questions 17-19 dealt with the theme of being or not being nervous in exams i.e. having tension in exams. Question 17 dealt with exams in general and question 18 was focused on exams on

speaking skills as shown in Figure 5. Thus, it was possible to compare tension in exams in general and in exams on speaking skills. The rating scales were from 0 to 10 in which 0 was *not at all* and 10 was *really a lot*. The average value of being nervous in exams in general was 3,55 and the median was 3. The standard deviation (SD) was 2,71. Conversely, the average value of being nervous in exams on speaking skills was almost a double: 6,52. Moreover, the median was over twice as high being 7. The SD was 2,85. Thus it can be concluded that exams on speaking skills should be even more carefully planned than exams on other language skills since only one respondent (3%) had marked tension being 10 *really a lot* in exams in general and seven (21%) had chosen it in speaking skills exams. 45% of the respondents, i.e. 15 respondents, had chosen 0-2 when asked if being nervous in exams in general, and conversely, 42% of the respondents, i.e. 14 respondents, had chosen 8-10 when asked about being nervous in speaking skills exams. Thus, the difference is clear.
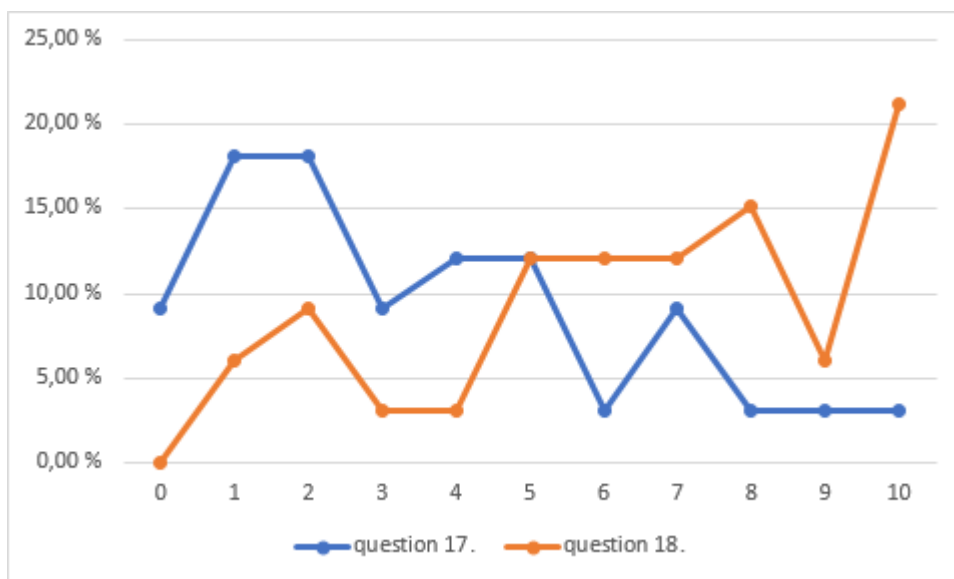


Figure 5. Comparison of the answers in question 17 'Are you nervous in exams **in general**?' and question 18 'Are you nervous in exams on **speaking skills**?'.

Question 19 asked about the effect the use of mobile phone applications could have on the respondents' possibility of being nervous in exams on speaking skills and it had four options to choose from (see Figure 6). The first two options had the clarifying 'Why'-questions. Using mobile phone applications could be seen a strength or weakness. All in all, 55% of respondents answered that using mobile phone applications would reduce tension in exams on speaking skills in English. 33% chose the answer 'I cannot say.'. 12% answered 'it would not have any effect at all'. None of the respondents chose the option 'it would increase it'. The answers to

the clarifying 'Why'-question was divided into two main categories: 1) no face-to-face situation (69%) and 2) you can try again (13%). 18% of the answers were not categorized. The reasons why a 1) no face-to-face situation would reduce tension were for example the following: *You wouldn't need to be in front of your teacher and your classmates (Ei tarvitsisi olla opettajan ja luokkalaisten edessä); you would not need to be directly in front of the person who is evaluating you (ei tarvitsisi olla suoraan arvioijan edessä)* and *a videosystem wouldn't necessarily be as exciting because it doesn't have to be in real time (jokin videosysteemi ei välttämättä olisi niin jönnittävä koska sen ei tarvitse olla reaaliajassa).* The reason for 2) you can try again, was that you can try again if you fail. Examples of non-categorized answers were for example, *it is more relaxed (se on rennompaa)* and *you could practise beforehand (voisi harjoitella etukäteen).*

As it can be seen in Figure 6 there were differences between men's and women's answers. Women were a bit more certain that it would reduce tension (57%) than men (50%) were. There were more men (50%) than women (26%) who were not able to say if it would reduce or increase tension or have any effect at all. While 17% of women thought it would not have any effect at all, none of the men chose this option.
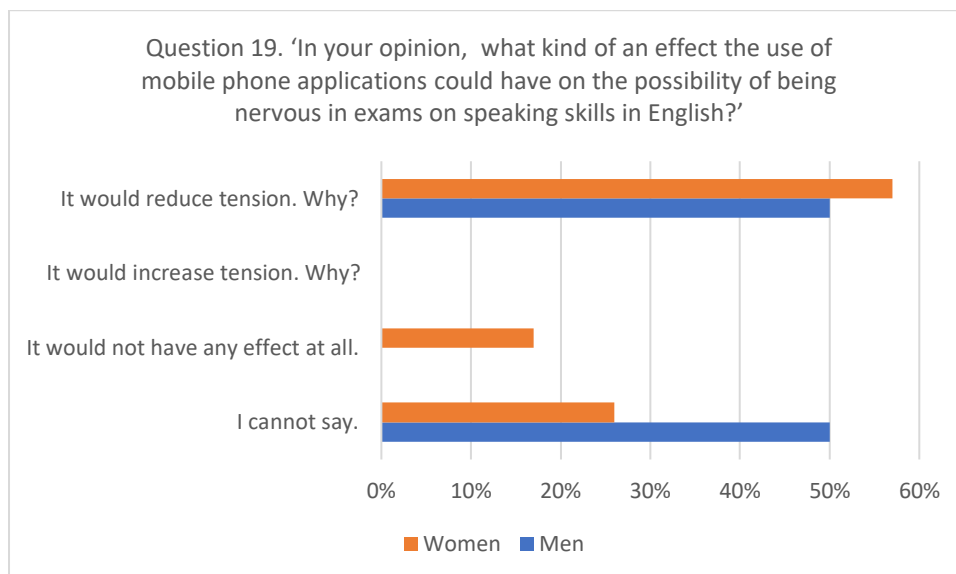


Figure 6. Respondents' opinions of the effect the use of mobile phone applications could have on the possibility of being nervous in exams on speaking skills in English. Answers compared between women and men.

4.3.3 The need for training

When analyzing answers for questions 20 and 21, it was easy to notice that students think that the need for training is higher for teachers (43%) than for themselves (18%) (see Figure 7). Question 20 concerned about students' need for training when using mobile phone applications in exams and question 21 teachers' need for training. 12% of the respondents answered that students do not need training whereas none of the respondents answered that teachers do not need training. 43% answered that training may be needed for students and 21% answered 'maybe' when asked about teachers. There were quite a few who did not know ('en osaa sanoa'): 27% when asked about students and 36% when asked about teachers.
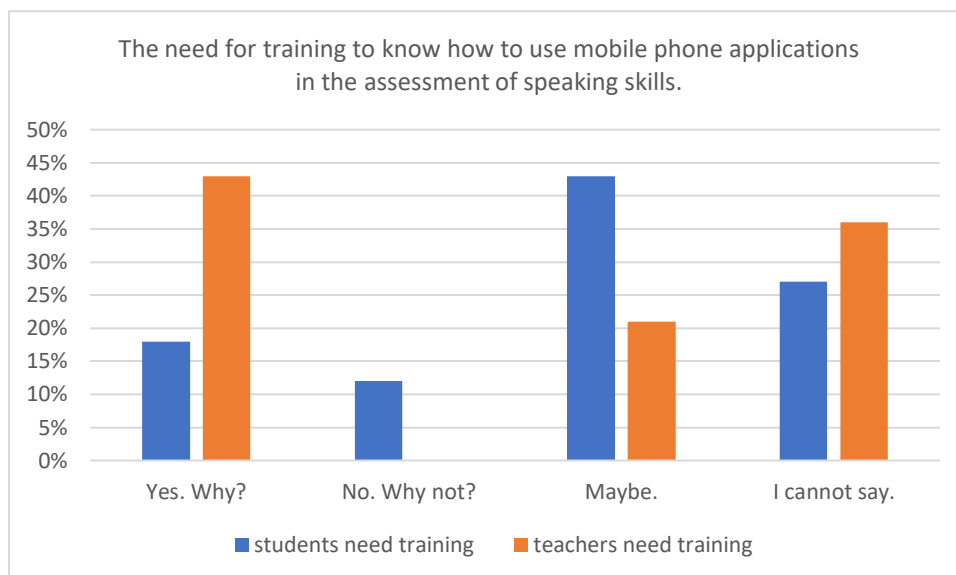


Figure 7 Comparing answers between the Question 20 'In your opinion, would students need training when using mobile phone applications in exams?' and the question 21 'In your opinion, would teachers need training when using mobile phone applications in exams?'.

In both questions 20 and 21 there were a clarifying question 'Why?' or 'Why not?' in the closed-ended items 'Yes.' and 'No.'(see Figure 7). When asked about students, the answers to this 'Why?' question in the 'Yes' answer were categorized in three main categories: 1) students are not familiar with the applications (60%), 2) stress (20%) and 3) usefulness/utility (20% =1 respondent). In the first category of 1) students are not familiar with the applications, the respondents underlined the two situations of a) if students are not familiar with the applications used in exams or b) if students have not used them. For example, one respondent answered that *If you don't know how to use an application, it is stressful (Jos sovellusta ei osaa käyttää, se on*

*stressaavaa).* Another respondent answered *It could be handy (Voisi olla kätevä).* There were two different kind of answers to the 'Why?' question in the 'No' answer: *You'll manage without it (Pärjeepi ilimanniin)* and *It's better orally* (*Suullisesti parenpi).*

The answers to the clarifying 'Why?' question when asked about teachers and the answer being 'Yes.' for training for teachers were categorized in two main categories: 1) they do not know how to use/well enough (45%) and 2) so that the teachers could know how to give advice to students (45%). There was also one empty answer, a line (-). For example, answers in the first main category were the following: *They do not even know how to use computers (Ne ei osaa käyttää edes tietokoneita)* and *Many teachers can use applications only poorly or insufficiently (Monet opettajat osaavat käyttää sovelluksia huonosti tai puutteellisesti)* Examples of the second main category were the following: *If a student needs help from the teacher (Jos oppilas tarvitsee opettajan apua); It is good if the teacher knows how to give advice (Opettajan on hyvä osata neuvoa)* and *they should know how to advise students to use applications but normally it goes the other way round (heidän pitäisi osata neuvoa oppilaita sovellusten käytössä mutta yleensä se menee toisin pain).* As we can see, some of the students found it important that their teachers would know how to help them when using new applications in the assessment of speaking skills.

# 5 Discussion

In this chapter, I will present the main results and at the end of it I will reflect on how I could have done different choices in my questionnaire after having the results. I will start by presenting the task types, implementation and smartphone applications respondents mentioned to be their preferred ones. Then the suggested smartphone applications in the assessment of speaking skills will be presented and after that there will be the SWOT analysis on the respondents' perceptions of using mobile phone applications. After that I will discuss test usefulness when using smartphone applications in the assessment of speaking skills and finally, there will be a critical evaluation of the research process.

The CEFR presents two main task categories for the assessment of the spoken production (Council of Europe 2018: 68-73): sustained monologue and addressing audiences and their subcategories of which some are same as the preferred task types the respondents mentioned in the questionnaire. The results show (see Table 2) that the type of the tasks most of the respondents (28%) preferred was *discussions.* Moreover, *reading aloud* and *a speech/a*

*presentation* were mentioned often (both 15%). As implementation the most preferred method was *in pairs* (22%). Furthermore, *in front of the class* (12%) and *with their teacher* (9%) were mentioned. As a conclusion it could be said that students would like to have discussions in pairs or maybe with their teacher and speeches or presentations in front of the class or alone with their teacher. Some students prefer the face-to-face situations from being in front of the class and some students do not want to have face-to-face exams and prefer for example to record at home and then send the recording to their teacher. Face-to-face situations compared to recordings could be an interesting topic for further research for the previous reasons. Derwing and Munro (2015: 124) mention audio-video software, such as *Skype* to be used in interactions between speakers and it would be interesting to study if the students who do not like traditional face-to-face situations could cope with using mobile phone applications like *Whatsapp*, *Skype, Zoom* or *Teams* in O365 in which there is a kind of face-to-face situation but not being physically in the same place. As Canale and Swain (1980) have argued verbal and non-verbal skills are a part of strategic competence and thus an important part of interaction in face-to-face situations. Mobile phone applications do not exclude the possibility to have face-to-face situations since students can be face-to-face when using applications when working in pairs or in groups. They can also give peer feedback and learn in the process from each other.

As a response for the RQ2 the analysis told me the alternatives of the smartphone applications which were the most top-rated among the respondents. These applications were recording applications (76%), the applications to make movies (61%) and role and avatar applications (39%). As a conclusion the students are ready to try various smartphone applications in the assessment of speaking skills even though there still are students who would not like to use recording, drama or role/avatar applications.

The analysis offered applications as a response to RQ3 (see Table 3). Question 14 asked about the applications respondents already use and if they could be used in the assessment of speaking skills in their opinion. This appeared to be a difficult item to the respondents to answer to. In questions 10-12 only five respondents suggested applications which were *Snapchat, camera, Gacha life* and *Bittmoji* but they may not be practical and useful smartphone applications to be used in the assessment of speaking skills. The reasons for the difficulty to answer these open-ended items might be the closed ended items in which suitable applications were already mentioned or the fact that the respondents did not know how to combine their knowledge of applications and the formal assessment of speaking skills. As Cochrane (2015: 138) has mentioned it is unclear "whether students are unaware of, or unwilling to use, the educational and productivity functions of their smartphones". None of the students mentioned

*Skype*, *Zoom* or *Teams* in O365 as possible smartphone applications to be used. Maybe because it is not always easy to remember and perceive that audio-video software used in CALL can also be used in MALL as applications.

Overall, the results told me the various perceptions the students had on using mobile phone applications in the formal assessment of speaking skills, as shown in Table 7, in which all the parts of SWOT are concluded. As strengths and weaknesses (see also Table 4 for more detailed information) the respondents mentioned issues which were easy to be defined as strengths or weaknesses but there were also some issues on which it was not easy to know which ones they had meant to be (Table 7). Repetition was mentioned as a strength, and indeed repetition and feedback are important when wanting to improve pronunciation (Kjellin 2002, as quoted by Kuronen 2017: 59-72). When using mobile phone applications students can repeat their performance several times and then listen to it. They can use self-assessment and learn from watching their own videos and listening to their pronunciation. Kuronen (2017: 59-72) emphasizes also perception as way to improve pronunciation and this can also be put into practice with applications. Assessment can be seen as a learning process.

Table 7. Answers to the RP presented in a form of a SWOT chart separating strengths and opportunities from weaknesses and threats. (questions 13, 15-22)

| **Strengths and Opportunities** | **Weaknesses and Threats** |
|---|---|
| the tension would be reduced/vanished | technical problems: the quality of the voice > could produce tension/stress |
| creativity | technical know-how: the difficulty to use these applications |
| new possibilities | > therefore, training needed for both students and teachers |
| diversity/variety | |
| easiness <br> • of making, processing, sending <br> • to use these applications <br> • to be yourself | the problems of concentration and focusing |
| Possible to repeat the performance <br> • fulfilment of students and thus the evaluation/assessment will be easier, more accurate and more diverse/versatile | the possibility to cheat* |
| **Strengths or weaknesses – can be interpreted as both depending on the case/student** | |
| *can be recorded several times/in parts* | *can be recorded several times/in parts* <br> *=*?* |
| *not face-to-face situations* | *not face-to-face situations* |

I will then discuss the opportunities of the applications which were categorized into those having significance for students or for teachers (Table 5). The respondents perceived quite easily the use of applications as an opportunity to reduce or even to vanish the tension. Thus, the importance of smartphone applications should be acknowledged and used in the assessment of speaking skills to reduce tension since tension can affect the results of the students. In the worst case, tension can prevent students to do exams in speaking skills. 55% of the respondents thought that using mobile phone applications would reduce the tension and none of the respondents thought it would increase it (question 19). This is a clear result to be acknowledged when planning exams.

Diversity is one of the opportunities that applications can give to the formal assessment of speaking skills and this was mentioned in the respondents' answers. Fučeková (2018) mentions that mobile phone applications are a powerful educational tool which will prove to be useful to both teachers and students. It was nice to see that students were able to consider applications as a powerful tool also for their teachers in the assessment of speaking skills because of the possibility to repeat the performance and thus have accurate evaluations. This could also be seen a part of test usefulness, as reliability (Bachman and Palmer 1996).

Finally, the threats of the SWOT analysis were categorized in technical know-how and psychological problems (see Table 6 for more detailed information). 18% of the respondents thought that students need training to know how to use applications and 43% that teachers need it. When the intention is to make assessment situations less stressful, it is important to consider the adequate training both for students and teachers. Abugohar et al. (2019) emphasized the importance of training in their study since teachers' perceptions of using smartphone applications in teaching speaking skills were positive but they were used less if there was not proper training to know how to use them effectively. The second main category of threats in my study were psychological problems and as that kind of problems the possibility to cheat was mentioned. Moreover, in my study the problems of concentrating were clearly a threat and also Cochrane (2015) mentioned smartphones as a distraction that keeps students from achieving success in their studies since students use them for maintaining social contacts and as gaming devices. For example, it is easy to imagine that a student could be disturbed if he/she was making a video as a test in speaking skills in Whatsapp and then received messages from friends.

I will next discuss test usefulness and its six components (Bachman and Palmer 1996: 17-40) when using smartphone applications in the assessment of speaking skills. Reliability can be achieved when using applications since they can provide consistency in the measurement,

but the tasks should be carefully designed to achieve it. Construct validity would be achieved if for example easily measured parts of prosody, such as pitch, would like to be assessed. But as we have seen, speaking skills consists of various issues and for example when measuring pragmatic competences such as cultural skills, it would be challenging. But we need to remember, that it might be even more challenging without the help of applications. Authenticity could be achieved since tasks in TLU contexts could be planned, for example a *Zoom* interaction with a native. If applications were used in matriculation exams in the future, then also impact could be achieved. The last of the six components of test usefulness, practicality, is maybe the easiest one to state since exams can be made at home and there is no need for a teacher to be present or there is no need to reserve any peaceful classroom to have exams.

I will now discuss the critical evaluation of the research process. Since the sample was only 33 respondents, in some of the optional open-ended items there were just a few answers and thus it was difficult to create categories and in some of the categories there were only isolated answers. Thus, it is possible that there could have been more categories of answers if there had been more respondents. Moreover, it was important to ask reasons for students' choices. All open-ended clarification items were not obligatory though for the reason being that if there had been too many obligatory items the respondents would have finished the questionnaire even less frequently than they did now. Furthermore, even though there were clarifying why-questions in closed-ended items, the results in these were not a success. But in some questions, there should have been obligatory clarification items. For example, in question 13 it would have been better if it had asked the respondents to explain if their answer was supposed to be seen a strength or a weakness because of the ambivalent answers (see Table 17). Moreover, in question 16 there should have been a clarification question which should have asked for example 'Explain in what kind of situations do you think the threat/threats you mentioned, would be seen as threats?'. Furthermore, a clarifying question of how to use the proposed applications in the assessment of speaking skills would have been useful to have (see Table 3) since it was not clear how the respondents meant the applications should be used. Conversely, the conclusions on task types and implementation (see Table 2) would have been clearer if there had been a closed-ended item in which various possibilities for both the task type and the implementation had been given and the respondents had made pairs of them. But an advantage of having an open-ended item was that it gave answers which might have been lost if only closed ended items and ready-made suggestions would have been used.

Mackey and Gass (2005) suggested several methods which could have been useful in the present study. It would have been interesting to conduct interviews with the respondents to

have deeper knowledge about the ambivalent or incomplete answers and to use in the questionnaire also oral answers which could have been recorded since they would have been a good choice for example for those who have limited literacy. For example, there was one respondent who had another L1 than Finnish who might have profited these methods. Moreover, they suggested that the questionnaire could have been piloted among the research population but because of the time limit it was not piloted among them.

## 6 Conclusion

The aim of the present study was to investigate the students' perceptions of using mobile phone applications in the formal assessment of speaking skills. SWOT analysis was used to divide the perceptions. Based on the results teachers could find practical and efficient ways to assess speaking skills using smartphone applications and thus offer their students more versatile methods to do it. Moreover, the perceptions of students of using mobile phone applications in the formal assessment of speaking skills can be taken into consideration.

When asked about previous exams in speaking skills, only 24% of the respondents answered having had them in lower comprehensive school and 46% answered having had them in upper comprehensive school. The result was higher in upper secondary school: 76% (see Table 1). It seems like teachers could be needing new easy to use tools in the formal assessment of speaking skills and mobile phone applications can provide multiple options to them. As discussed in Chapter 5, practicality of the six components of test usefulness is of high value when using applications. As already mentioned in the previous studies on MALL (Abugohar et al. 2019, Cochrane 2015, Fučeková 2018), the smartphone applications offer us new possibilities in the field of teaching and learning language skills in the EFL context in the 21st century and thus they can also offer us new possibilities in the assessment of speaking skills.

In the fields of speaking skills, assessment and MALL there are several ideas for the future research. One of them is to study how many teachers have in equal proportions tests on different language 'skills' or their combinations – reading comprehension, writing, grammar, vocabulary, listening comprehension, speaking and interaction, and of which of these 'skills' Wilma's average grade is composed before deciding the final grades for students. Moreover, it would be interesting and useful to study what kind of possibilities there are for the assessment of speaking skills on electronic platforms, for example on Otava's and Sanomapro's electronic platforms for exams. Furthermore, teachers' and students' perceptions of assessing orally could

be studied and if it would be possible to give easily feedback in oral form in Wilma instead of written form. In addition to CALL and MALL there is Robot assisted language learning (RALL) which could also facilitate teachers work and offer various new methods. The future of the assessment of speaking skills could be studied and more precisely the perceptions of teachers and students of MALL, CALL and RALL.

It is always important to remember that students are customers and their opinions matter, also in the assessment. All in all, the majority of the students on upper secondary school were willing to use mobile phone applications in the formal assessment of speaking skills and they saw it as a possibility but that there was also uncertainty if the applications worked well and if teachers and students knew how to use them properly. The results and the reproduction of the present study can be useful since teachers can have tools to work with and students can influence the ways of the formal assessment of speaking skills. As a result of my study, a new service – a course or training – during which smartphone applications that could be used in the assessment of speaking skills would be introduced to teachers and students. There could be recording, drama, and role/avatar applications in which drama as a teaching, learning and assessing method could be used. Students could be in a role and thus their character/avatar would protect them if they are afraid of speaking English or are nervous in exams. As Fučeková (2018: 173) states "mobile applications represent a powerful and popular educational tool which is worth attention of teachers, researchers, and academics worldwide, and further investigation into this matter is necessary, and it will certainly prove useful to both teachers and students."

# Bibliography

Abugohar, M.A., Yunus, K. and Rashid, R. A. 2019. Smartphone applications as a Teaching Technique for Enhancing Tertiary Learners' Speaking Skills: Perceptions and Practices. *iJET International Journal of Emerging Technologies in Learning.* Vol 14 No 9 (2019). https://doi.org/10.3991/ijet.v14i09.10375

Ahola, S. (2017). Suullisen kielitaidon arviointi. In E. Tergujeff and M. Kautonen (eds), *Suullinen kielitaito. Opi, opeta ja arvioi*. Otava, 153-168.

Bachman, L. F. and Palmer A.S. 1996. *Language Testing in Practice.* Oxford University Press. China.

Canale, M & Swain, M. 1980. Theoretical bases of communicative approaches to second language teaching and testing. *Applied linguistics*, 1, 1-47.

Cochrane, R. 2015. Activities and reflection for influencing beliefs about learning with smartphones. In F. Helm, L. Bradley, M. Guarda, & S. Thouësny (Eds), Critical CALL – Proceedings of the 2015 EUROCALL Conference, Padova, Italy (pp. 138-143). Dublin: Research-publishing.net. http://dx.doi.org/10.14705/ rpnet.2015.000323

Council of Europe (2020). Common European Framework for Reference for languages: Qualitative aspects of spoken language use - Table 3 (CEFR 3.3): Common Reference levels. https://www.coe.int/en/web/common-european-framework-reference-languages/table-3-cefr-3.3-common-reference-levels-qualitative-aspects-of-spoken-language-use

Council of Europe (2018). Common European Framework of reference for languages: learning, teaching, assessment. Companion volume with new descriptors. (Language Policy Programme, Education Policy Division, Education Department). https://rm.coe.int/cefr-companion-volume-with-new-descriptors-2018/1680787989

Derwing, T. M. & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. Amsterdam, Netherlands; Philadelphia, Pennsylvania: John Benjamins Publishing Company.

Dörnyei, Z. & Taguchi, T. (2010). *Questionnaires in second language research: Construction, administration, and processing* (2nd ed.). New York; London: Routledge.

Fučeková, M. 2018. Developing English Skills by means of Mobile Applications. *Information Technologies and Learning Tools*, Vol 66, No 4: 173-185.

Johnson, K. 2013. An introduction to foreign language learning and teaching. (2nd ed.) New York. Routledge.

Keurulainen, H. (2013). Pelisääntöjä arviointipäätösten tekemistä varten. In E. Räisänen (ed), *Oppimisen arvioinnin kontekstit ja käytännöt. Raportit ja selvitykset.* Opetushallitus, 3: 37-60.

Kuronen, M. (2017). Ääntämisen foneettiset oppimiskohteet. In E. Tergujeff and M. Kautonen (eds), *Suullinen kielitaito. Opi, opeta ja arvioi*. Otava, 59-74.

Lukion opetussuunnitelman perusteet 2015. Finnish National Board of Education (online). https://www.oph.fi/sites/default/files/documents/172124_lukion_opetussuunnitelman_peruste et_2015.pdf. (pv, kk 2020 –esim. 26, January 2020).

Lukion opetussuunnitelman perusteet 2019. Finnish National Board of Education (online). https://www.oph.fi/sites/default/files/documents/lukion_opetussuunnitelman_perusteet_2019. pdf. (…. Kts. Edeltä).

Luoma, S. 2004. *Assessing Speaking.* Cambridge University Press. New York.

Mackey, A. & Gass, S. M. (2005). *Second language research: Methodology and design.* Mahwah, NJ: Lawrence Erlbaum.

Piccardo, E. 2016. Common European Framework of Reference for Languages: learning, teaching, assessment. Phonological scale revision process report. (Language Policy). Council of Europe. https://rm.coe.int/phonological-scale-revision-process-report-cefr/168073fff9

Tergujeff, E. (2017). Ääntäminen - keskeiset kielikohtaiset oppimiskohteet. Englanti. In E. Tergujeff and M. Kautonen (eds), *Suullinen kielitaito. Opi, opeta ja arvioi.* Otava, 170-171.

Tergujeff, E., Heinonen, H., Ilola, M., Salo, O.-P. and Kara H. (2017). Suullisen kielitaidon opetus käytännössä. In E. Tergujeff and M. Kautonen (eds), *Suullinen kielitaito. Opi, opeta ja arvioi.* Otava, 96-117.

Tergujeff, E. and Kautonen, M. (2017). Suullinen kielitaito ja sen tärkeys. In E. Tergujeff and M. Kautonen (eds), *Suullinen kielitaito. Opi, opeta ja arvioi.* Otava, 12-21.

Tolmie, A., Muijs, D. & McAteer, E. (2011). *Quantitative methods in educational and social research using SPSS.* Maidenhead: Open University Press.

Appendix: Questionnaire (in Finnish)

**Tietosuojailmoitus**

**Kännykkäsovellusten käyttö englannin suullisen kielitaidon arvioinnissa**

Tietoa tutkimuksesta

Olen englannin maisteriopiskelija Katja Jääskeläinen ja toteutan tämän kyselyn osana maisterintutkielmaani eli graduani Jyväskylän yliopiston kieli- ja viestintätieteiden laitoksella. Tutkimukseni tavoitteena on kartoittaa, **miten lukion oppilaana haluaisit tulla arvioiduksi englannin kielen suullisessa kielitaidossa**.

Sinulla on nyt mahdollisuus tuoda mielipiteesi esiin tästä asiasta. Kyselyn pääpaino on **kännykkäsovellusten hyödyntämisessä suullisen kielitaidon muodollisessa arvioinnissa**. Muodollisella arvioinnilla tarkoitan virallisia koetilanteita, joissa suullista kielitaitoasi arvioidaan arvosanoin (sanallinen arviointi tai numero). Tämä kysely sisältää myös yleisluonteisia kysymyksiä suullisen kielitaidon arvioinnista.

Tutkimus tuottaa tietoa siitä, millaisia kännykkäsovelluksia lukiolaisten mielestä olisi mukava käyttää suullista kielitaitoa arvioitaessa, miten suullista kielitaitoasi on tähän mennessä arvioitu eri kouluasteilla ja millaisia suullisia kokeita itse toivoisit pidettävän.

Vastaamalla tähän kyselyyn annat suostumuksesi vastausten käyttämiseen tutkimuksessani. Halutessasi voit kieltäytyä osallistumasta kyselyyn ilman minkäänlaisia seuraamuksia. Voit perua osallistumisesi myös myöhemmin, jolloin vastauksiasi ei käytetä tutkimuksessani. Opettajasi eivät saa kyselyn tuloksia käyttöönsä, mutta kyselystä tekemäni analyysit tulevat osaksi tutkimustuloksiani. Valmista tutkielmaa voi sen valmistuttua käydä lukemassa Jyväskylän yliopiston kirjaston sähköisessä JYX-palvelussa.

Kyselyn vastaukset analysoidaan määrällisesti niin, että yksittäisiä vastaajia ei voi tunnistaa. Tutkimusaineistoa säilytetään Jyväskylän yliopiston tutkimusaineiston käsittelyä koskevien tietoturvakäytänteiden mukaisesti ilman vastaajien tunnistetietoja ja siten, että kukaan ulkopuolinen ei pääse aineistoon käsiksi.

Kyselyyn vastaaminen kestää noin 10-15 minuuttia.
Kyselyssä ei ole oikeita tai vääriä vastauksia. Sinun mielipiteesi on tärkeintä.
**Kiitos paljon arvokkaista vastauksistasi!**


Ystävällisin terveisin,

Englannin maisteriopiskelija
*Katja Jääskeläinen*
*kaejaask@student.jyu.fi*

**Kyselylomake**

*pakollinen vastattava

TAUSTATIEDOT

1. **Vuosikurssisi***
- 1. vuosikurssi
- 2. vuosikurssi
- 3. vuosikurssi

2. **Sukupuolesi***
- nainen
- mies
- muu
- en halua sanoa

3. **Äidinkielesi***
- suomi
- Muu kuin suomi. Mikä?

4. **Viimeisimmän englannin kurssisi arvosana lukiossa?***
- 4
- 5-6
- 7-8
- 9-10
- Minulla ei ole vielä yhdenkään englannin kurssin arvosanaa lukiosta. Yläkoulun päättöarvosana oli ___

YLEISIÄ KYSYMYKSIÄ ENGLANNIN SUULLISISTA KOKEISTA

5. **Oliko sinulla alakoulussa englannin suullisen kielitaidon kokeita, joista sait arvosanan?***
- Kyllä. Millaisia?
- Ei.
- En muista./En osaa sanoa.

6. **Oliko sinulla yläkoulussa englannin suullisen kielitaidon kokeita, joista sait arvosanan?***
- Kyllä. Millaisia?

- Ei.
- En muista./En osaa sanoa.

7. **Onko sinulla ollut lukiossa englannin suullisen kielitaidon kokeita, joista sait arvosanan?***
- Kyllä. Millaisia?
- Ei.
- En muista./En osaa sanoa.

8. **Jos saisit itse päättää, millaisia englannin suullisia kokeita Sinulle pidettäisiin, niin millaisia ne olisivat?***

9. **Miten haluaisit, että Sinua arvioitaisiin englannin suullisissa kokeissa?** *Esimerkiksi* **numeroin, sanallisin arvioin tai muulla tavoin.**

ENGLANNIN SUULLISEN KIELITAIDON ARVIOINTI KÄNNYKKÄSOVELLUKSIN

10. **Mitä kännykän ÄÄNITTÄMISSOVELLUKSIA haluaisit käyttää englannin suullisissa kokeissa?**

*Esimerkiksi* **englannin suullisessa kokeessa, jossa olisi opettajan antama englanninkielinen sanalista tai teksti, joka luettaisiin ääneen ja lähetettäisiin opettajalle.**

**Voit valita yhden tai useamman vaihtoehdon.***

- puheen äänittämis-ohjelmia (esim. Androidissa nimellä Ääninauhuri ja iPhonessa nimellä Sanelin)
- Whatsapp
- Jokin muu/jotkin muut äänittämissovellukset. Mikä/mitkä?
- En haluaisi käyttää kännykän äänittämissovelluksia. Miksi et?

11. **Mitä kännykän DRAAMASOVELLUKSIA haluaisit käyttää englannin suullisissa kokeissa?**

*Esimerkiksi* **englannin suullisessa kokeessa, jossa olisi opettaja antaisi tehtäväksesi kuvitteellisen kauppareissun, jonka tekisitte pareittain tai pienissä ryhmissä.**

**Voit valita yhden tai useamman vaihtoehdon.***

- elokuvien teko-sovelluksia (esim. iMovie…)
- nukketeatteri-sovelluksia (esim. Puppet Pals…)
- 'suunnittele ja äänitä oma kertomuksesi'-sovelluksia (esim. Lipa Theater…)
- Jokin muu/jotkin muut draamasovellukset. Mikä/mitkä?
- En haluaisi käyttää kännykän draamasovelluksia. Miksi et?

**12. Mitä kännykän ROOLIHAHMO- tai AVATAR-SOVELLUKSIA haluaisit käyttää englannin suullisissa kokeissa?**

*Esimerkiksi* englannin suullisessa kokeessa, jossa opettaja antaisi tehtäväksesi kuvitella olevasi vanhus, joka kertoisi elämästään.

**Voit valita yhden tai useamman vaihtoehdon.***

- Avatar-sovelluksia (esim. Morfo…)
- nukketeatteri-sovelluksia (esim. Puppet Pals…)
- 'suunnittele ja äänitä oma kertomuksesi'-sovelluksia (esim. Lipa Theater…)
- Jokin muu/jotkin muut draamasovellukset. Mikä/mitkä?
- En haluaisi käyttää kännykän roolihahmo- tai avatar-sovelluksia. Miksi et?


**13. Kerro, mitä hyviä ja huonoja puolia on ÄÄNITTÄMIS-, DRAAMA-, ROOLIHAHMO- tai AVATAR-kännykkäsovelluksissa, joita olet jo käyttänyt?**


**14. Mitä kännykkäsovelluksia, joita jo käytät, voisit käyttää englannin suullisten kokeiden tekemiseen?***


**15. Mitä mahdollisuuksia kännykkäsovellukset voisivat mielestäsi tuoda englannin suullisen kielitaidon arviointiin?**


**16. Mitä haittoja kännykkäsovellukset voisivat mielestäsi tuoda englannin suullisen kielitaidon arviointiin?**


YLEISIÄ KYSYMYKSIÄ KOETILANTEISTA


17. **Jännitätkö yleisesti koetilanteita? Siirrä liukukytkin itsellesi sopivaan paikkaan. Kännykällä vastatessasi napauta viivaa siltä kohdalta, jonka haluat olevan vastauksesi**.

0=en lainkaan ja 10=todella paljon*


18. **Jännitätkö yleisesti koetilanteita? Siirrä liukukytkin itsellesi sopivaan paikkaan. Kännykällä vastatessasi napauta viivaa siltä kohdalta, jonka haluat olevan vastauksesi.**

0=en lainkaan ja 10=todella paljon*

**19. Millä tavalla Sinun mielestäsi kännykkäsovellusten käyttö voisi vaikuttaa mahdolliseen jännittämiseen englannin suullisissa kokeissa?***

- Vähentäisi jännittämistä. Miksi?
- Pahentaisi jännittämistä. Miksi?
- Ei vaikuttaisi mitenkään.
- En osaa sanoa.

**20. Tarvitsisiko mielestäsi oppilaille olla kännykkäsovellus-koulutusta koekäyttö-tarkoituksia varten?***

- Kyllä. Miksi?
- Ei. Miksi?
- Ehkä.
- En osaa sanoa.

**21. Tarvitsisiko mielestäsi opettajille olla kännykkäsovellus-koulutusta koekäyttö-tarkoituksia varten?***

- Kyllä. Miksi?
- Ei. Miksi?
- Ehkä.
- En osaa sanoa.

Kiitos paljon arvokkaista vastauksistasi!
Have a nice day :)

W

Kysely luotu Webropolilla
Klikkaa tästä ja lue lisää